

# EXCEPTIONAL SETS FOR DIOPHANTINE INEQUALITIES

SCOTT T. PARSELL AND TREVOR D. WOOLEY

**ABSTRACT.** We apply Freeman's variant of the Davenport-Heilbronn method to investigate the exceptional set of real numbers not close to some value of a given real diagonal form at an integral argument. Under appropriate conditions, we show that the exceptional set in the interval  $[-N, N]$  has measure  $O(N^{1-\delta})$ , for a positive number  $\delta$ .

## 1. INTRODUCTION

A variant of the classical circle method introduced by Davenport and Heilbronn [17] permits the investigation of the value distribution of indefinite real diagonal forms at integral points. Let  $k \in \mathbb{N}$  and let  $\lambda_1, \dots, \lambda_s$  be non-zero real numbers, not all in rational ratio, and not all of the same sign when  $k$  is even. Then the Davenport-Heilbronn method establishes the existence of a natural number  $\mathfrak{s}(k)$  having the property that, whenever  $s \geq \mathfrak{s}(k)$ , then for all  $0 < \tau \leq 1$  and  $\mu \in \mathbb{R}$ , there exist infinitely many integral solutions  $\mathbf{x}$  of the inequality

$$|\lambda_1 x_1^k + \cdots + \lambda_s x_s^k - \mu| < \tau. \quad (1.1)$$

When  $\mu = 0$  and  $\mathfrak{s}(k) = 2^k + 1$ , a result of this type is described, for example, in [35, Theorem 11.1]. For smaller values of  $s$ , available technology may limit accessible conclusions to analogues involving some sort of averaging over the real number  $\mu$ . Quantitative estimates in this direction are ultimately connected to the savings achievable on a suitably defined set of minor arcs, and so even the pivotal reorganization of the Davenport-Heilbronn method introduced by Freeman [19] apparently limits such bounds to be better than trivial only by the narrowest of margins. In this paper we demonstrate that, in a wide set of circumstances, this barrier may be unequivocally broken. Indeed, we show that the Diophantine inequality (1.1) is satisfied for all real numbers  $\mu \in [-N, N]$  with the possible exception of a set having measure  $O(N^{1-\Delta})$ , for a positive number  $\Delta$ .

In order to advance further, we must formalise the above ideas and introduce notation with which to describe our conclusions. Let  $\mathcal{Z}_{s,k}(N, M) = \mathcal{Z}_{s,k}^\tau(N, M; \boldsymbol{\lambda})$  denote the set of real numbers  $\mu \in [N, N+M]$  for which the inequality (1.1) has no integral solution. On writing

$$F(\mathbf{x}) = \lambda_1 x_1^k + \cdots + \lambda_s x_s^k \quad (1.2)$$

and

$$\mathfrak{B} = \bigcap_{\mathbf{x} \in \mathbb{Z}^s} (-\infty, F(\mathbf{x}) - \tau] \cup [F(\mathbf{x}) + \tau, \infty),$$

---

2010 *Mathematics Subject Classification.* 11D75.

STP was supported by National Security Agency Grants H98230-08-1-0070 and H98230-11-1-0190 and TDW by a Royal Society Wolfson Research Merit Award.

one sees that  $\mathcal{Z}_{s,k}(N, M) = \mathfrak{B} \cap [N, N + M]$ . This shows that  $\mathcal{Z}_{s,k}(N, M)$  is a closed subset of  $\mathbb{R}$  and in particular that it is measurable. We may therefore write

$$Z_{s,k}(N, M) = \text{meas}(\mathcal{Z}_{s,k}(N, M)).$$

Although the solubility of (1.1) for all  $\mu$  requires an indefiniteness hypothesis, our results on exceptional sets apply equally well to situations in which the form  $F(\mathbf{x})$  is definite. The proofs do require slightly more care in the definite case to ensure compatibility between the size of  $\mu$  and the ranges of the variables, but this is easily arranged by summing over dyadic intervals. We concentrate on the case where  $F$  has at least one positive coefficient exceeding 2 and leave to the reader the necessary sign changes and re-scaling required to formulate the general case.

Our results are concisely introduced in some generality by reference to available minor arc estimates for exponential sums. With this end in mind we record some additional notation. When  $P$  and  $R$  are positive real numbers, write

$$\mathcal{A}(P, R) = \{n \in [1, P] \cap \mathbb{Z} : p|n \Rightarrow p \leq R\}.$$

Here and throughout, we employ the letter  $p$  to denote a prime number, and we assume that  $R$  is a sufficiently small positive power of  $P$ . In addition, we put

$$g(\alpha) = \sum_{x \in \mathcal{A}(P, R)} e(\alpha x^k), \quad (1.3)$$

where as usual  $e(z)$  denotes  $e^{2\pi iz}$ . We next define a standard set of major and minor arcs. When  $q \in \mathbb{N}$  and  $a \in \mathbb{Z}$ , we take  $\mathfrak{N}(q, a)$  to be the set of real numbers  $\alpha$  for which  $|q\alpha - a| \leq (2k)^{-1}P^{1-k}$ . We write  $\mathfrak{N}$  for the union of the intervals  $\mathfrak{N}(q, a)$  over all co-prime integers  $a$  and  $q$  with  $1 \leq q \leq (2k)^{-1}P$ , and we put  $\mathfrak{n} = \mathbb{R} \setminus \mathfrak{N}$ . Finally, we say that the pair  $(s, \sigma)$  forms a *smooth accessible pair for  $k$*  when  $s \geq 2k$ , and for some positive number  $\omega$ , whenever  $\varepsilon > 0$ , then

$$\int_{\mathfrak{n} \cap [0, 1]} |g(\alpha)|^s d\alpha \ll P^{s-k-\omega} \quad \text{and} \quad \sup_{\alpha \in \mathfrak{n}} |g(\alpha)| \ll P^{1-\sigma+\varepsilon}. \quad (1.4)$$

In §6 we establish a bound on the measure of the exceptional set  $\mathcal{Z}_{s,k}(N, M)$ . Here and elsewhere, we adopt the convention that whenever  $\varepsilon$  appears in a statement, then we implicitly assert that the statement holds for all  $\varepsilon > 0$ . Moreover, implicit constants in Vinogradov's notation may depend on  $s$ ,  $k$ ,  $\lambda$  and  $\tau$ .

**Theorem 1.1.** *Suppose that  $k \geq 3$  and that  $(s_0, \sigma_0)$  forms a smooth accessible pair for  $k$ . Let  $s$  and  $t$  be non-negative integers with*

$$s \geq \max\{2k + 3, 17, \frac{1}{2}s_0\}.$$

*Then whenever  $M \geq N^{(1-1/k)^t}$ , there exists a positive number  $\Delta$  such that*

$$Z_{s+t,k}(N, M) \ll MN^{-\Delta-(2s-s_0)\sigma_0/k}. \quad (1.5)$$

Subject to the hypotheses of the statement of this theorem, the methods discussed in [46] permit the proof of the bound  $Z_{s,k}(N, M) \ll 1$  when  $s \geq s_0$ . Indeed, when the form  $\lambda_1 x_1^k + \dots + \lambda_s x_s^k$  is indefinite the latter bound may be replaced by the definitive statement that  $Z_{s,k}(N, M) = 0$ . In §§6 and 7 we apply the work of [12, 37, 38, 39, 41] to provide a refined explicit version of Theorem 1.1.

**Theorem 1.2.** Suppose that  $4 \leq k \leq 20$ , and that  $s_0 = s_0(k)$ ,  $u_0 = u_0(k)$  and  $\sigma = \sigma(k)$  are as given in Table 1. Then whenever  $s$  and  $t$  are non-negative integers with  $s \geq \frac{1}{2}(s_0 + u_0)$ , and  $M \geq N^{(1-1/k)^t}$ , there exists a positive number  $\Delta$  such that

$$Z_{s+t,k}(N, M) \ll MN^{-\Delta-(2s-s_0)\sigma/k}.$$

Moreover, when  $s \geq \frac{1}{2}s_0$ , and  $M \geq N^{(1-1/k)^t}$ , then for some  $\Delta > 0$  one has

$$Z_{s+t,k}(N, M) \ll MN^{-\Delta}.$$

The same conclusions hold for larger values of  $k$  on setting  $u_0(k) = 0$ ,

$$s_0(k) = k(\log k + \log \log k + 2 + o(1)) \quad \text{and} \quad \sigma(k)^{-1} = k(\log k + O(\log \log k)).$$

$k$	4	5	6	7	8	9	10	11	12
$s_0(k)$	12	18	25	33	42	50	59	67	76
$u_0(k)$	4	6	7	0	0	0	0	0	0
$\sigma(k)^{-1}$	8	16	32	58	70	83	95	108	120

  

$k$	13	14	15	16	17	18	19	20
$s_0(k)$	84	92	100	109	117	125	134	142
$u_0(k)$	0	0	0	0	0	0	0	0
$\sigma(k)^{-1}$	133	146	158	171	184	197	210	223

Table 1: Parameters for Theorem 1.2

Write  $\mathcal{Z}_{s,k}^\tau(N; \boldsymbol{\lambda})$  for the set of real numbers  $\mu \in [-N, N]$  for which the Diophantine inequality (1.1) has no solution, and let  $Z_{s,k}(N)$  denote its measure. Then by applying Theorem 1.1 with  $t = 0$  and summing over dyadic intervals, one finds that there is a positive number  $\Delta$  such that, whenever

$$s \geq \max\{2k + 3, 17, \frac{1}{2}s_0(k)\},$$

then  $Z_{s,k}(N) = O(N^{1-\Delta})$ . This confirms the estimate advertised in our opening paragraph. Furthermore, if  $t \sim \beta k \log k$ , where  $0 < \beta \leq \frac{1}{2}$ , then one finds that Theorem 1.1 applies with intervals of length  $M \asymp N^\gamma$ , where  $\gamma = k^{-\beta+o(1)}$ . Hence if  $s_0(k) \geq k \log k$ , then there are integers  $s < s_0(k)$  for which the estimate (1.5) holds with  $M$  of approximate order  $N^{1/\sqrt{k}}$ .

In the special case  $k = 3$ , additional control may be exercised over exponential sum estimates, and this permits several refinements over the conclusion of Theorem 1.1. We illustrate such ideas in §8 by establishing the following theorem.

**Theorem 1.3.** Let  $\xi$  be a real number satisfying  $\xi^{-1} > 2556 + 48\sqrt{2833}$ . Then

$$Z_{4,3}(N) = o(N), \quad Z_{5,3}(N) \ll N^{3/4-\xi}, \quad \text{and} \quad Z_{6,3}(N) \ll N^{1/2-2\xi}.$$

Analogous conclusions for Waring's problem are sharper, most notably for sums of four cubes (see [6, 25, 42, 43]). Experts will find the explanation in the absence of a  $p$ -adic iteration restricted to minor arcs in the context of Diophantine inequalities.

We now consider asymptotic formulae subject to the restriction that  $0 < \tau \leq 1$ . Denote by  $\mathcal{N}_{s,k}^\tau(P; \boldsymbol{\lambda}, \mu)$  the number of integral solutions of the inequality (1.1) with  $\mathbf{x} \in [1, P]^s$ , and note our earlier assumption that no coefficient  $\lambda_i$  is zero.

We put  $\nu_i = |\lambda_i|$  and  $\sigma_i = \lambda_i/\nu_i$ , and then define  $\mathfrak{V}(\theta)$  to be the set of  $\mathbf{v} \in [0, \nu_1] \times \cdots \times [0, \nu_{s-1}]$  satisfying the condition

$$\sigma_s(\theta - \sigma_1 v_1 - \cdots - \sigma_{s-1} v_{s-1}) \in [0, \nu_s].$$

Finally, we write

$$\Omega_{s,k}(\boldsymbol{\lambda}, \theta) = k^{-s} |\lambda_1 \dots \lambda_s|^{-1/k} C_{s,k}(\boldsymbol{\lambda}, \theta),$$

where

$$C_{s,k}(\boldsymbol{\lambda}, \theta) = \int_{\mathfrak{V}(\theta)} (v_1 \dots v_{s-1})^{1/k-1} (\sigma_s(\theta - \sigma_1 v_1 - \cdots - \sigma_{s-1} v_{s-1}))^{1/k-1} d\mathbf{v}.$$

When  $k \geq 3$  and  $s \geq k+1$ , it follows from a heuristic application of the Davenport-Heilbronn method that there is a function  $L(P)$  tending to infinity such that

$$\mathcal{N}_{s,k}^\tau(P; \boldsymbol{\lambda}, \mu) = 2\tau \Omega_{s,k}(\boldsymbol{\lambda}, \mu P^{-k}) P^{s-k} + O(P^{s-k} L(P)^{-1}). \quad (1.6)$$

We note that  $1 \ll \Omega_{s,k}(\boldsymbol{\lambda}, \theta) \ll 1$  provided only that  $\text{meas}(\mathfrak{V}(\theta)) \gg 1$ , and in such circumstances the relation (1.6) constitutes an honest asymptotic formula.

Next, let  $z$  be a positive parameter, and consider a positive function  $\psi(z)$  growing sufficiently slowly in terms of  $z$ . We denote by  $\tilde{\mathcal{Z}}_{s,k}^\tau(N; \psi; \boldsymbol{\lambda})$  the set of real numbers  $\mu \in (N/2, N]$  for which one has

$$|\mathcal{N}_{s,k}^\tau(N^{1/k}; \boldsymbol{\lambda}, \mu) - 2\tau \Omega_{s,k}(\boldsymbol{\lambda}, \mu/N) N^{s/k-1}| > N^{s/k-1} \psi(N)^{-1}. \quad (1.7)$$

Recall the notation introduced in (1.2) and write  $\chi(\mu; a, b)$  for the indicator function of the interval  $(a, b)$ , so that

$$\chi(\mu; a, b) = \begin{cases} 1, & \text{when } a < \mu < b, \\ 0, & \text{when } \mu \leq a \text{ or } \mu \geq b. \end{cases}$$

Then we see that the counting function  $\mathcal{N}_{s,k}^\tau(\mu) = \mathcal{N}_{s,k}^\tau(N^{1/k}; \boldsymbol{\lambda}, \mu)$  can be defined by means of the relation

$$\mathcal{N}_{s,k}^\tau(\mu) = \sum_{1 \leq x_1, \dots, x_s \leq N^{1/k}} \chi(\mu; F(\mathbf{x}) - \tau, F(\mathbf{x}) + \tau).$$

Our earlier assumption that one at least of the coefficients  $\lambda_i$  is positive and exceeds 2 implies that  $\text{meas}(\mathfrak{V}(\mu/N)) > 0$ , and hence  $\Omega = \Omega_{s,k}(\boldsymbol{\lambda}, \mu/N) > 0$ . We therefore find that  $\mathcal{N}_{s,k}^\tau(\mu) - 2\tau \Omega N^{s/k-1}$  is a measurable function of  $\mu$ , and hence the set  $\tilde{\mathcal{Z}}_{s,k}^\tau(N; \psi; \boldsymbol{\lambda})$  is measurable. We write

$$\tilde{Z}_{s,k}(N) = \text{meas}(\tilde{\mathcal{Z}}_{s,k}^\tau(N; \psi; \boldsymbol{\lambda})).$$

As in our earlier discussion of the counting function  $Z_{s,k}(N)$ , we introduce some notation with which to discuss minor arc estimates for classical Weyl sums. Write

$$f(\alpha) = \sum_{1 \leq x \leq P} e(\alpha x^k).$$

We say that the triple  $(s, \sigma, U)$  forms an *accessible triple for  $k$*  when  $s \geq 2k$ , the function  $U(P)$  increases monotonically to infinity, and one has

$$\int_{\mathfrak{n} \cap [0,1]} |f(\alpha)|^s d\alpha \ll P^{s-k} U(P)^{-1} \quad \text{and} \quad \sup_{\alpha \in \mathfrak{n}} |f(\alpha)| \ll P^{1-\sigma+\varepsilon}. \quad (1.8)$$

In §3 we investigate the measure of the exceptional set  $\tilde{Z}_{s,k}(N)$ .

**Theorem 1.4.** *Suppose that  $k \geq 3$  and that  $(s_1, \sigma_1, U)$  forms an accessible triple for  $k$ . Then whenever*

$$s \geq \max\{k + 2, \frac{1}{2}s_1\}, \quad \sigma < \sigma_1,$$

*and  $\psi(N)$  grows sufficiently slowly in terms of  $s$ ,  $\sigma$ ,  $k$ ,  $\lambda$ , and  $\tau$ , one has*

$$\tilde{Z}_{s,k}(N) \ll N^{1-(2s-s_1)\sigma/k} U(N^{1/k})^{\varepsilon-1}.$$

Although Theorem 1.4 does not address estimates for  $\tilde{Z}_{4,3}(N)$ , a fairly pedestrian approach yields the estimate contained in the following theorem.

**Theorem 1.5.** *Whenever  $\psi(N)$  grows sufficiently slowly in terms of  $\lambda$  and  $\tau$ , one has  $\tilde{Z}_{4,3}(N) = o(N)$ .*

Subject to the hypotheses of the statement of Theorem 1.4, the methods discussed in [46] may be applied to show that  $\tilde{Z}_{s,k}(N) \ll 1$  when  $s \geq s_1$ . The best accessible triples for smaller values of  $k$  stem from work of Boklan [3] and Vaughan [31], [33]. Thus, for a suitable positive number  $\gamma$ , one has the accessible triples

$$(s_1, \sigma_1, U(N)) = (2^k, 2^{1-k}, (\log N)^\gamma) \quad (3 \leq k \leq 5).$$

For larger values of  $k$  the picture has recently been transformed by developments stemming from the second author's efficient congruencing approach to Vinogradov's mean value theorem (see [48], [49], [50]). In particular, on defining the exponent  $\sigma^*(k)$  by means of the relation

$$1/\sigma^*(k) = \begin{cases} 2^{k-1}, & \text{when } k = 6, 7, \\ 2k(k-2), & \text{when } k \geq 8, \end{cases}$$

it follows that for a suitable positive number  $\gamma$  one has the accessible triples

$$(s_1, \sigma_1, U(N)) = (2k^2 - 2k - 8, \sigma^*(k), N^\gamma) \quad (k \geq 6)$$

and

$$(s_1, \sigma_1, U(N)) = (2k^2 - 2, \sigma^*(k), N^{1/k-\varepsilon}) \quad (k \geq 6).$$

Weaker conclusions of similar type could be extracted from the earlier work of Boklan [4], Ford [18], Heath-Brown [22], and Wooley [40]. For a comprehensive discussion of the various ingredients, see [46, §7].

When  $s$  is relatively close to  $s_1$ , one can obtain conclusions sharper than those available from Theorem 1.4 by including some of the excess variables in a mean value together with an integral over the exceptional set. For example, by adapting the "slim" technology of Wooley [47], along with the refinements of Kawada and Wooley [25], [26], in §4 we derive the following refinement for smaller exponents.

**Theorem 1.6.** *Suppose that  $\psi(N)$  grows sufficiently slowly in terms of  $\lambda$  and  $\tau$ . Then with the values of  $k$ ,  $s$  and  $\beta$  from the following table, it follows that for every  $\varepsilon > 0$  one has  $\tilde{Z}_{s,k}(N) \ll N^{\beta+\varepsilon}$ .*

$k$	3	4	4	4	5	5	5	5	5	5	5
$s$	7	13	14	15	25	26	27	28	29	30	31
$\beta$	$\frac{1}{3}$	$\frac{5}{8}$	$\frac{1}{2}$	$\frac{7}{16}$	$\frac{3}{4}$	$\frac{7}{10}$	$\frac{13}{20}$	$\frac{3}{5}$	$\frac{23}{40}$	$\frac{11}{20}$	$\frac{3}{8}$

Table 2: Exponents for slim exceptional sets

A direct application of Theorem 1.4 would yield weaker estimates. Thus, for example, whereas Theorem 1.4 shows that  $\tilde{Z}_{7,3}(N) \ll N^{1/2+\varepsilon}$ , one sees from Theorem 1.6 that  $\tilde{Z}_{7,3}(N) \ll N^{1/3+\varepsilon}$ .

At the cost of requiring an extra variable, in §5 we provide a short-interval analogue of Theorem 1.4. Here we write  $\tilde{\mathcal{Z}}_{s,k}^\tau(N, M; \psi; \boldsymbol{\lambda})$  for the set of real numbers  $\mu \in [N, N+M]$  for which (1.7) holds, and  $\tilde{Z}_{s,k}(N, M)$  for its measure.

**Theorem 1.7.** *Suppose that  $k \geq 3$  and that  $(s_1, \sigma_1, U)$  forms an accessible triple for  $k$ . Then whenever*

$$s \geq \max\{k+1, \frac{1}{2}s_1\}, \quad \sigma < \sigma_1, \quad M \geq N^{1-1/k},$$

and  $\psi(N)$  grows sufficiently slowly in terms of  $s$ ,  $\sigma$ ,  $k$ ,  $\boldsymbol{\lambda}$ , and  $\tau$ , one has

$$\tilde{Z}_{s+1,k}(N, M) \ll MN^{-(2s-s_1)\sigma/k}U(N^{1/k})^{\varepsilon-1}.$$

When  $s$  is somewhat smaller than is required to bound the exceptional set  $\mathcal{Z}_{s,k}^\tau(N; \boldsymbol{\lambda})$  successfully, we can instead aim for non-trivial lower bounds for the measure of the set of  $\mu \in [-N, N]$  for which the inequality (1.1) does have a solution. We let  $Y_{s,k}(N)$  denote the measure of the set

$$\mathcal{Y}_{s,k}^\tau(N; \boldsymbol{\lambda}) = [-N, N] \setminus \mathcal{Z}_{s,k}^\tau(N; \boldsymbol{\lambda}).$$

As in the analogous questions related to Waring's problem, the lower bounds for  $Y_{s,k}(N)$  which we derive in §9 depend upon suitable upper bounds for mean values of exponential sums. We say that the exponent  $\Delta = \Delta_{s,k}$  is *admissible* if for each  $\varepsilon > 0$ , whenever  $R \leq P^\eta$  and  $\eta$  is sufficiently small, one has

$$\int_0^1 |g(\alpha)|^{2s} d\alpha \ll P^{2s-k+\Delta+\varepsilon}. \quad (1.9)$$

**Theorem 1.8.** *If  $\Delta = \Delta_{s,k}$  is admissible, then for every  $\varepsilon > 0$  one has*

$$Y_{s,k}(N) \gg \tau^2 N^{1-\Delta/k-\varepsilon}.$$

In particular, one has  $Y_{3,3}(N) \gg \tau^2 N^{\gamma-\varepsilon}$ , where  $\gamma = (166 - \sqrt{2833})/123 > 11/12$ .

We note that Theorem 1.8 delivers a non-trivial conclusion even when  $\tau$  is an explicit function of  $N$ , provided only that  $\tau^{-1}$  is somewhat smaller than  $(N^{1-\Delta/k})^{1/2}$ .

Finally, we consider the approximation of real numbers by linear combinations of two primes, a topic which may be viewed as an analogue of the binary Goldbach problem. Suppose that  $\lambda_1$  and  $\lambda_2$  are real numbers with  $\lambda_1/\lambda_2$  irrational. Let  $\mathcal{Z}^*(X; \boldsymbol{\lambda}, \tau)$  denote the set of real numbers  $\mu \in [0, X]$  for which the inequality

$$|\lambda_1 p_1 + \lambda_2 p_2 - \mu| < \tau$$

has no solution in prime numbers  $p_1, p_2$ . In §10 we derive a conclusion free of the spacing condition on  $\mu$  present in earlier work of Brüdern, Cook and Perelli [7] and the first author [28].

**Theorem 1.9.** *One has  $\text{meas}(\mathcal{Z}^*(X; \boldsymbol{\lambda}, \tau)) = o(X)$ .*

It would appear that a quantitative version of this result is currently inaccessible, although in the special case where  $\lambda_1/\lambda_2$  is algebraic, such a conclusion has recently been obtained by Brüdern, Kawada and Wooley [10]. In this situation, an explicit bound on the failures of the asymptotic formula is achieved by [10, Theorem 1.5].

The results of this paper are motivated by analogous considerations in Waring's problem, in which for suitable values of  $s$  one seeks to represent all sufficiently large integers  $n$  in the form  $x_1^k + \cdots + x_s^k = n$ . When  $s$  fails to be large enough to establish such a conclusion, one may instead try to bound the number of "exceptional" integers  $n$  that fail to admit such a representation. In particular, if the Hardy-Littlewood method produces representations for all large  $n$  whenever  $s \geq \mathcal{G}(k)$ , then one can typically show that the number of exceptional integers  $n \leq X$  is  $o(X)$  whenever  $s \geq \frac{1}{2}\mathcal{G}(k)$ . There has been considerable recent work aimed at refining our understanding of these exceptional sets. Although quantitative bounds are ultimately connected to the savings that one can achieve on a suitably defined set of minor arcs, the power and flexibility of the methods has recently been enhanced by technology in which an exponential sum over the set of exceptions is used to better exploit extra variables (see [8], [9], [15], [25], [26], [27], [44], [45] and [47]).

Historically, the situation for Diophantine inequalities has been somewhat different. For inequalities of the shape (1.1), the method of Davenport and Heilbronn requires one to restrict to a possibly sparse sequence of box sizes defined in terms of the continued fraction convergents to some ratio  $\lambda_i/\lambda_j$ . Such a restriction on the box size offers little hope of analysing the set of exceptional  $\mu$  for which (1.1) has no integer solution. However, work of Freeman [19], [20], inspired by the methods of Bentkus and Götze [2] and refined by the second author [46], has changed the perspective. As a result of these innovations, one is now able to obtain asymptotics for the number of solutions in all sufficiently large boxes, albeit with inexplicit error terms arising from the minor and trivial arcs. Moreover, it transpires that quantitative upper bounds on the measure of the exceptional set are accessible by employing Hardy-Littlewood dissections with respect to various  $\lambda_i\alpha$ , where  $\alpha$  is the variable of integration. This idea has been useful in previous work on inequalities (see for example the proof of [29, Lemma 2.3], and the amplification procedure of [46]) when attempting to obtain optimal estimates for mixed mean values. The points for which none of the  $\lambda_i\alpha$  satisfy a classical minor arc condition are handled using the Bentkus-Götze-Freeman technology, which suffices to show that the contribution from such  $\alpha$  is negligible in comparison to the main term and allows one to exploit the quantitative savings available elsewhere. The method can also be applied to analyse failures of the expected asymptotic formula and to the situation where  $\mu$  ranges over a short interval. In contrast to some previous approaches (see for example [7], [28]) in which the  $\mu$  under consideration were assumed to satisfy an unnatural spacing condition, we make no such assumption and instead integrate directly over the exceptional set.

A word is in order concerning the history of the present paper. An early preprint of this work existed in mid-2007, with exceptional sets containing power savings in a revision prepared following an Oberwolfach meeting a year later. This work had some influence on the paper on thin sequences [10] joint between Brüdern, Kawada and the second author (see [11] for subsequent developments). With an eye to the delays created by workload obstructions, we have taken the opportunity in this final version of incorporating the very recent developments pertaining to classical Weyl sums stemming from work of the second author on Vinogradov's mean value theorem (see [48], [49] and [50]).

## 2. THE KERNEL OF THE ANALYSIS

In this section we introduce the elements of the Davenport-Heilbronn method key to our discussion, as well as some basic estimates required during the course of our argument. When  $P$  is a large positive number and  $1 \leq Q \leq P$  we write

$$f(\alpha; Q, P) = \sum_{Q < x \leq P} e(\alpha x^k). \quad (2.1)$$

Given two suitable positive functions  $S(P)$  and  $T(P) \leq S(P)$  tending to infinity, we define the major arc by

$$\mathfrak{M} = \{\alpha \in \mathbb{R} : |\alpha| \leq S(P)P^{-k}\}.$$

We further write

$$\mathfrak{m} = \{\alpha \in \mathbb{R} : S(P)P^{-k} < |\alpha| \leq T(P)\} \quad \text{and} \quad \mathfrak{t} = \{\alpha \in \mathbb{R} : |\alpha| > T(P)\}$$

for the minor and trivial arcs, and set  $L(P) = \max\{1, \log T(P)\}$ . We apply a dissection of this same basic shape for all of our applications, although the specific form of the functions  $S$ ,  $T$ , and  $L$  may change (compare for instance the definitions in [46, §3 and §8]). For the moment, it suffices to note that this set-up ensures that the major, minor, and trivial arcs give the contributions normally expected in Freeman's version of the Davenport-Heilbronn method when the number of variables is sufficiently large.

We next recall the upper and lower bound kernels  $K_{\pm}(\alpha)$  defined by Freeman [20]. Write  $\delta = \tau L(P)^{-1}$  and put

$$K_{\pm}(\alpha) = \frac{\sin(\pi\delta\alpha) \sin(\pi(2\tau \pm \delta)\alpha)}{\pi^2 \delta \alpha^2}. \quad (2.2)$$

Also, write  $U_a(t)$  for the indicator function of the interval  $(-a, a)$ . Then from [20, Lemma 1] and its proof, one finds that

$$U_{\tau-\delta}(t) \leq \int_{-\infty}^{\infty} e(\alpha t) K_{-}(\alpha) d\alpha \leq U_{\tau}(t) \quad (2.3)$$

and

$$U_{\tau}(t) \leq \int_{-\infty}^{\infty} e(\alpha t) K_{+}(\alpha) d\alpha \leq U_{\tau+\delta}(t). \quad (2.4)$$

Moreover, one has the bound

$$K_{\pm}(\alpha) \ll \min\{\tau, |\alpha|^{-1}, \delta^{-1} \alpha^{-2}\}. \quad (2.5)$$

In order to facilitate analytic manipulations, it is convenient to work with positive kernels. On recalling (2.2), we find that when necessary the kernel  $K_{\pm}(\alpha)$  may be decomposed by means of the relation

$$|K_{\pm}(\alpha)|^2 = K_1(\alpha)K_2^{\pm}(\alpha), \quad (2.6)$$

where

$$K_1(\alpha) = \left( \frac{\sin(\pi\delta\alpha)}{\pi\delta\alpha} \right)^2 \leq \min\{1, (\pi\delta\alpha)^{-2}\} \quad (2.7)$$

and

$$K_2^{\pm}(\alpha) = \left( \frac{\sin(\pi(2\tau \pm \delta)\alpha)}{\pi\alpha} \right)^2 \leq \min\{5\tau^2, (\pi\alpha)^{-2}\}. \quad (2.8)$$

We note that by making a change of variable in [1, Lemma 14.1], one finds that the Fourier transform of the kernel  $K_1$  satisfies

$$\widehat{K}_1(t) = \int_{-\infty}^{\infty} e(\alpha t) K_1(\alpha) d\alpha = \delta^{-1} \max\{0, 1 - \delta^{-1}|t|\} \leq \delta^{-1} U_{\delta}(t). \quad (2.9)$$

Next we introduce exponential integrals with which to encode the sets of real numbers that we seek to test for unrepresentations, or occasionally for representations. When  $\mathcal{Z}$  is any measurable subset of  $\mathbb{R}$  and  $\eta_{\mu}$  is any complex-valued function of  $\mu$ , put

$$H_{\eta, \mathcal{Z}}(\alpha) = \int_{\mathcal{Z}} \eta_{\mu} e(-\alpha\mu) d\mu. \quad (2.10)$$

The following mean value estimates play a critical rôle in our arguments.

**Lemma 2.1.** *When  $\mathcal{Z}$  is a set of finite measure  $Z$ , and  $|\eta_{\mu}| = 1$  for  $\mu \in \mathcal{Z}$ , then*

$$\int_{-\infty}^{\infty} |H_{\eta, \mathcal{Z}}(\alpha)|^2 K_1(\alpha) d\alpha \leq 2Z.$$

*Proof.* In view of (2.7) and (2.10), it follows from Fubini's Theorem that the mean value in question may be written as

$$I = \int_{\mathcal{Z}^2} \overline{\eta_{\mu}} \eta_{\nu} \int_{-\infty}^{\infty} e(\alpha(\mu - \nu)) K_1(\alpha) d\alpha d\mu d\nu = \int_{\mathcal{Z}^2} \overline{\eta_{\mu}} \eta_{\nu} \widehat{K}_1(\mu - \nu) d\mu d\nu.$$

Moreover, the upper bound (2.9) shows that for every fixed  $\nu$  one has  $\widehat{K}_1(\mu - \nu) = 0$  unless  $\nu - \delta < \mu < \nu + \delta$ . Consequently,

$$I \leq \int_{\mathcal{Z}^2} \widehat{K}_1(\mu - \nu) d\mu d\nu \leq \delta^{-1} \int_{\mathcal{Z}} \int_{\nu-\delta}^{\nu+\delta} d\mu d\nu = 2Z,$$

and the proof of the lemma is complete.  $\square$

When  $c$  is a positive constant we write  $f^{(c)}(\alpha)$  for the exponential sum  $f(\alpha; cP, P)$  defined in (2.1). The following lemma allows us to handle exceptional sets in short intervals and plays a key rôle in the proof of Theorem 1.7.

**Lemma 2.2.** *Let  $N$  be a large real number, and write  $P = N^{1/k}$ . In addition, let  $c$  and  $\lambda$  be non-zero real numbers with  $c > 0$ . Suppose that  $M \leq |\lambda|(k-1)(cP)^{k-1}$ .*

Then whenever  $\mathcal{Z}$  is a subset of  $[N, N + M]$  having measure  $Z$ , and  $|\eta_\mu| = 1$  for all  $\mu \in \mathcal{Z}$ , one has

$$\int_{-\infty}^{\infty} |f^{(c)}(\lambda\alpha)H_{\eta,\mathcal{Z}}(\alpha)|^2 K_1(\alpha) d\alpha \leq 2PZ.$$

*Proof.* As in the proof of Lemma 2.1, the mean value under consideration may be written as

$$I = \int_{\mathcal{Z}^2} \bar{\eta}_\mu \eta_\nu \int_{-\infty}^{\infty} |f^{(c)}(\lambda\alpha)|^2 e(\alpha(\mu - \nu)) K_1(\alpha) d\alpha d\mu d\nu.$$

In view of the relation (2.9), one has

$$I \leq \int_{\mathcal{Z}^2} \delta^{-1} \mathcal{W}(\mu, \nu) d\mu d\nu, \quad (2.11)$$

where  $\mathcal{W}(\mu, \nu)$  denotes the number of solutions of the inequality

$$|\lambda(x^k - y^k) + \mu - \nu| < \delta,$$

with  $cP < x, y \leq P$ . Our hypothesis concerning the size of  $M$  ensures that whenever  $\mu, \nu \in \mathcal{Z}$ , one has

$$|\mu - \nu| \leq M \leq |\lambda|(k-1)(cP)^{k-1}.$$

Then if  $cP < x, y \leq P$  and  $x \neq y$ , one has

$$\begin{aligned} |\lambda(x^k - y^k)| &= |\lambda(x-y)(x^{k-1} + x^{k-2}y + \dots + y^{k-1})| > |\lambda|k(cP)^{k-1} \\ &> |\lambda|(k-1)(cP)^{k-1} + \delta \geq |\mu - \nu| + \delta. \end{aligned}$$

We therefore deduce that  $x = y$  for every solution counted by  $\mathcal{W}(\mu, \nu)$ , and hence that  $\mathcal{W}(\mu, \nu) \leq P$ . Moreover, for every fixed  $\nu$  one has  $\mathcal{W}(\mu, \nu) = 0$  unless  $\nu - \delta < \mu < \nu + \delta$ . Thus we infer from (2.11) that

$$I \leq \delta^{-1} P \int_{\mathcal{Z}} \int_{\nu-\delta}^{\nu+\delta} d\mu d\nu = 2PZ,$$

and the desired conclusion follows.  $\square$

We next record a general principle that allows us to extend mean value estimates over subsets of the unit interval to corresponding subsets of the whole real line in the presence of a suitably decaying kernel function.

**Lemma 2.3.** *Let  $X$  and  $t$  be non-negative real numbers, and let  $\lambda \in \mathbb{R}$ . Also, when  $\mathcal{A} \subseteq [1, P] \cap \mathbb{Z}$  and  $\mathfrak{B} \subseteq [0, 1)$  is measurable, let*

$$h(\alpha) = \sum_{x \in \mathcal{A}} e(\alpha x^k) \quad \text{and} \quad \mathfrak{B}_\lambda(X) = \{\alpha \in \mathbb{R} : |\alpha| \geq X \text{ and } \lambda\alpha \in \mathfrak{B} \pmod{1}\}.$$

*Then for any real-valued function  $K$  satisfying  $|K(\alpha)| \ll \min\{1, \alpha^{-2}\}$ , one has*

$$\int_{\mathfrak{B}_\lambda(X)} |h(\lambda\alpha)|^t K(\alpha) d\alpha \ll_\lambda (1+X)^{-1} \int_{\mathfrak{B}} |h(\alpha)|^t d\alpha.$$

*Proof.* After splitting into intervals of the form  $\mathfrak{B}_\lambda(X) \cap [n, n+1]$  and then making a change of variable, we obtain

$$\int_{\mathfrak{B}_\lambda(X)} |h(\lambda\alpha)|^t K(\alpha) d\alpha \ll |\lambda|^{-1} \sum_{n \geq X} (1+n)^{-2} \int_{\mathfrak{C}_\lambda(n)} |h(\alpha)|^t d\alpha,$$

where  $\mathfrak{C}_\lambda(n)$  denotes the set of  $\alpha \in \mathfrak{B} \pmod{1}$  with  $\lambda n \leq \alpha < \lambda(n+1)$ . Since  $\mathfrak{C}_\lambda(n)$  is contained in a union of at most  $1 + |\lambda|$  translates of  $\mathfrak{B}$  of the shape  $\mathfrak{B} + j$ , the conclusion of the lemma follows from the periodicity of  $h(\alpha)$  modulo 1.  $\square$

### 3. FAILURES OF THE ASYMPTOTIC FORMULA

We illustrate our methods first with the proof of Theorem 1.4, which concerns the frequency with which the anticipated asymptotic formula (1.6) fails for the function  $\mathcal{N}_{s,k}^\tau(P; \boldsymbol{\lambda}, \mu)$  counting the number of integral solutions of the inequality (1.1) with  $\mathbf{x} \in [1, P]^s$ . Observe that there is no loss of generality in supposing throughout that  $\lambda_1/\lambda_2 \notin \mathbb{Q}$ . Suppose that  $(s_1, \sigma_1, U)$  is an accessible triple for  $k$ , and put  $S(P) = \min\{U(P), (2k)^{-1}P\}$ . We begin by recording a minor arc estimate. Recall the definition (2.1) and write  $f_i(\alpha) = f(\lambda_i \alpha; 0, P)$ .

**Lemma 3.1.** *There exists a choice for the function  $T(P)$ , depending only on  $\lambda_1, \lambda_2$  and  $S(P)$ , with the property that*

$$\sup_{\alpha \in \mathfrak{m}} |f_1(\alpha) f_2(\alpha)| \ll P^2 T(P)^{-2^{-k-1}}.$$

*Proof.* In view of our definition of the minor arcs  $\mathfrak{m}$ , the desired conclusion is immediate from [46, Lemma 2.3].  $\square$

It follows from the conclusion of this lemma that for all  $\alpha \in \mathfrak{m}$ , one has

$$f_j(\alpha) \ll P T(P)^{-6^{-k}} \tag{3.1}$$

for at least one suffix  $j \in \{1, 2\}$ . We define  $\mathfrak{q}_j$  to be the set of real numbers  $\alpha \in \mathfrak{m}$  for which the upper bound (3.1) holds, and then put  $\mathfrak{p}_j = \mathfrak{q}_j \cup \mathfrak{t}$ , so that  $\mathfrak{m} \cup \mathfrak{t} \subseteq \mathfrak{p}_1 \cup \mathfrak{p}_2$ . In addition, when  $1 \leq i \leq s$ , we write  $\mathfrak{N}_i$  for the set of real numbers  $\alpha \in \mathfrak{m} \cup \mathfrak{t}$  for which  $\lambda_i \alpha \in \mathfrak{N}$ , and  $\mathfrak{n}_i = (\mathfrak{m} \cup \mathfrak{t}) \setminus \mathfrak{N}_i$ . In order to assist our discussion, we define

$$v_{ij} = \begin{cases} 1, & \text{when } i = j \text{ and } j \in \{1, 2\}, \\ 0, & \text{otherwise.} \end{cases}$$

Finally, when  $t > 0$ , the set  $\mathfrak{B}$  is measurable, and  $K$  is integrable, we write

$$\mathcal{M}_{i,t}(\mathfrak{B}; K) = \int_{\mathfrak{B}} |f_i(\alpha)|^t K(\alpha) d\alpha \quad (1 \leq i \leq s). \tag{3.2}$$

We first establish an estimate for the mean value introduced in (3.2) of modified major arc type.

**Lemma 3.2.** *Suppose that  $k \geq 3$ ,  $t > k+1$ ,  $1 \leq i \leq s$  and  $j \in \{1, 2\}$ . Then for any fixed  $\kappa > 0$ , one has*

$$\mathcal{M}_{i,t}(\mathfrak{N}_i \cap \mathfrak{p}_j; K_2^\pm) \ll_t P^{t-k} L(P)^{-\kappa s v_{ij}}$$

and

$$\mathcal{M}_{i,t}(\mathfrak{N}_i \cap \mathfrak{p}_j; |K_{\pm}|) \ll_t P^{t-k} L(P)^{1-\kappa s v_{ij}}.$$

*Proof.* We prove first that when  $t > k + 1$  and  $K(\alpha) \ll \min\{1, \alpha^{-2}\}$ , one has

$$\mathcal{M}_{i,t}(\mathfrak{N}_i \cap \mathfrak{p}_j; K) \ll_t P^{t-k} L(P)^{-\kappa s v_{ij}}. \quad (3.3)$$

We may suppose that  $t = k + 1 + 2\gamma$ , where  $\gamma > 0$ . We put  $u = k + 1 + \gamma$ . Then from the proof of Theorem 4.4 and Lemma 4.9 of [35], on using (4.13) of the latter in place of (4.14), we find that

$$\int_{\mathfrak{N}_i \cap [0,1)} |f_i(\alpha)|^u d\alpha \ll P^{u-k}.$$

Then by applying (3.1) in combination with Lemma 2.3, we have on the one hand

$$\begin{aligned} \int_{\mathfrak{N}_i \cap \mathfrak{q}_j} |f_i(\alpha)|^t |K(\alpha)| d\alpha &\ll \left( \sup_{\alpha \in \mathfrak{q}_j} |f_i(\alpha)| \right)^{\gamma} \int_{\mathfrak{N}_i \cap [0,1)} |f_i(\alpha)|^u d\alpha \\ &\ll \left( P T(P)^{-v_{ij} 6^{-k}} \right)^{\gamma} P^{u-k} \ll P^{t-k} L(P)^{-\kappa s v_{ij}}, \end{aligned}$$

whilst on the other

$$\begin{aligned} \int_{\mathfrak{N}_i \cap t} |f_i(\alpha)|^t |K(\alpha)| d\alpha &\ll f(0)^{\gamma} T(P)^{-1} \int_{\mathfrak{N}_i \cap [0,1)} |f_i(\alpha)|^u d\alpha \\ &\ll (P^{\gamma} L(P)^{-\kappa s}) P^{u-k} = P^{t-k} L(P)^{-\kappa s}. \end{aligned}$$

Our claimed bound (3.3) follows from (3.2) by combining these two upper bounds.

The respective conclusions of the lemma follow from (3.3) on noting first from (2.8) that  $K_2^{\pm}(\alpha) \ll \min\{1, \alpha^{-2}\}$ , and second from (2.5) that

$$K_{\pm}(\alpha) \ll \delta^{-1} \min\{1, \alpha^{-2}\} \ll L(P) \min\{1, \alpha^{-2}\}.$$

□

We turn next to a corresponding estimate of minor arc flavour.

**Lemma 3.3.** *Suppose that  $(s_1, \sigma_1, U)$  is an accessible triple for  $k$ , and that  $t$  and  $\sigma$  are real numbers with  $t \geq s_1$  and  $\sigma < \sigma_1$ . Then for  $1 \leq i \leq s$ , one has*

$$\mathcal{M}_{i,t}(\mathfrak{n}_i; K_2^{\pm}) \ll P^{t-k-(t-s_1)\sigma} U(P)^{-1}.$$

*Proof.* On noting (2.8), the desired estimate follows by inserting the estimates stemming from (1.8) into the conclusion of Lemma 2.3. □

We now embark upon the proof of Theorem 1.4. Adopt the hypotheses of the statement of the latter theorem, take  $N$  to be a positive number sufficiently large in terms of  $s, k, \boldsymbol{\lambda}$  and  $\tau$ , and put  $P = N^{1/k}$ . When  $\mu \in (N/2, N]$ , we define

$$R_{\pm}(\mu) = \int_{-\infty}^{\infty} \tilde{f}(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha,$$

in which we have written  $\tilde{f}(\alpha) = f_1(\alpha)f_2(\alpha)\cdots f_s(\alpha)$ . Then it follows from (2.3) and (2.4) that whenever  $\mu \in (N/2, N]$ , one has

$$R_{-}(\mu) \leq \mathcal{N}_{s,k}^{\tau}(N^{1/k}; \boldsymbol{\lambda}, \mu) \leq R_{+}(\mu). \quad (3.4)$$

In view of our assumption that  $\lambda_i > 2$  for some index  $i$ , the argument leading to [46, Lemma 6.1] shows that whenever  $\mu \in (N/2, N]$ , then

$$\int_{\mathfrak{M}} \tilde{f}(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha = 2\tau \Omega_{s,k}(\boldsymbol{\lambda}, \mu/N) P^{s-k} + O(P^{s-k} L(P)^{-1}).$$

Here we have made use of the implicit hypothesis that  $s \geq k+2$ . It follows that whenever  $\psi(N)$  grows slowly enough in terms of  $s, \sigma, k, \boldsymbol{\lambda}$  and  $\tau$ , and in particular, sufficiently slowly in terms of  $L(P)$ , then

$$\left| \int_{\mathfrak{M}} \tilde{f}(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha - 2\tau \Omega_{s,k}(\boldsymbol{\lambda}, \mu/N) N^{s/k-1} \right| < \frac{1}{2} N^{s/k-1} \psi(N)^{-1}. \quad (3.5)$$

Next, write  $\mathcal{Z} = \tilde{\mathcal{Z}}_{s,k}^{\tau}(N; \psi; \boldsymbol{\lambda})$ , put  $Z = \text{meas}(\mathcal{Z})$ , and consider an element  $\mu$  of  $\mathcal{Z}$ . Since  $\mathbb{R}$  is the disjoint union of  $\mathfrak{M}, \mathfrak{m}$  and  $\mathfrak{t}$ , a comparison of (1.7) and (3.5) leads from (3.4) to the conclusion that with  $K_{\mu} = K_+$  or  $K_{\mu} = K_-$ , one has

$$\left| \int_{\mathfrak{m} \cup \mathfrak{t}} \tilde{f}(\alpha) e(-\alpha\mu) K_{\mu}(\alpha) d\alpha \right| \geq \frac{1}{2} P^{s-k} \psi(N)^{-1}. \quad (3.6)$$

Using  $*$  to denote either  $+$  or  $-$ , denote by  $\mathcal{Z}^*$  the set of  $\mu \in \mathcal{Z}$  for which (3.6) holds with  $K_{\mu} = K_*$ , and write  $Z^* = \text{meas}(\mathcal{Z}^*)$ . Then it follows that  $Z \leq Z^+ + Z^-$ , so that for some choice of  $*$ , either  $+$  or  $-$ , one has  $Z \leq 2Z^*$ . We fix this choice henceforth. For each  $\mu \in \mathcal{Z}^*$ , we determine the complex number  $\eta_{\mu}$  by means of the relation

$$\left| \int_{\mathfrak{m} \cup \mathfrak{t}} \tilde{f}(\alpha) e(-\alpha\mu) K_*(\alpha) d\alpha \right| = \eta_{\mu} \int_{\mathfrak{m} \cup \mathfrak{t}} \tilde{f}(\alpha) e(-\alpha\mu) K_*(\alpha) d\alpha.$$

Recall the definition (2.10), write  $H(\alpha) = H_{\eta, \mathcal{Z}^*}(\alpha)$ , and note that  $|\eta_{\mu}| = 1$  for each  $\mu \in \mathcal{Z}^*$ . Then by integrating the relation (3.6) over the set  $\mathcal{Z}^*$ , we find that

$$\int_{\mathfrak{m} \cup \mathfrak{t}} \tilde{f}(\alpha) H(\alpha) K_*(\alpha) d\alpha \geq \frac{1}{2} P^{s-k} \psi(N)^{-1} \int_{\mathcal{Z}^*} d\mu.$$

For the sake of convenience, when  $\mathfrak{B}$  is measurable, we write

$$I(\mathfrak{B}) = \int_{\mathfrak{B}} |\tilde{f}(\alpha) H(\alpha) K_*(\alpha)| d\alpha. \quad (3.7)$$

In this notation, the last lower bound implies that

$$I(\mathfrak{p}_1) + I(\mathfrak{p}_2) \geq \frac{1}{4} P^{s-k} \psi(N)^{-1} Z. \quad (3.8)$$

In the final phase of our proof we obtain estimates for the integrals  $I(\mathfrak{p}_j)$ , thereby converting the lower bound (3.8) into an upper bound for  $Z$ . Let  $j$  be either 1 or 2. Then an application of Hölder's inequality leads from (3.7) to the bound

$$I(\mathfrak{p}_j) \leq \prod_{i=1}^s \left( \int_{\mathfrak{p}_j} |f_i(\alpha)^s H(\alpha) K_*(\alpha)| d\alpha \right)^{1/s}. \quad (3.9)$$

Consider an index  $i$  with  $1 \leq i \leq s$ . Since we may suppose that  $s \geq k + 2$ , we deduce from Lemma 3.2 that

$$\begin{aligned} \int_{\mathfrak{N}_i \cap \mathfrak{P}_j} |f_i(\alpha)^s H(\alpha) K_*(\alpha)| d\alpha &\leq H(0) \mathcal{M}_{i,s}(\mathfrak{N}_i \cap \mathfrak{P}_j; |K_*|) \\ &\ll Z P^{s-k} L(P)^{1-3sv_{ij}}. \end{aligned} \quad (3.10)$$

Next write

$$\mathcal{J} = \int_{-\infty}^{\infty} |H(\alpha)|^2 K_1(\alpha) d\alpha. \quad (3.11)$$

Then since we may suppose that  $s \geq \frac{1}{2}s_1$ , an application of Schwarz's inequality in combination with Lemmata 2.1 and 3.3 leads via (2.6) to the bound

$$\begin{aligned} \int_{\mathfrak{N}_i} |f_i(\alpha)^s H(\alpha) K_*(\alpha)| d\alpha &\leq \mathcal{J}^{1/2} \mathcal{M}_{i,2s}(\mathfrak{N}_i; K_2^*)^{1/2} \\ &\ll Z^{1/2} (P^{2s-2k} \Xi_{\varepsilon})^{1/2}, \end{aligned} \quad (3.12)$$

where we have written

$$\Xi_{\varepsilon} = P^{k-(2s-s_1)(\sigma_1-\varepsilon)} U(P)^{-1}. \quad (3.13)$$

On substituting (3.10) and (3.12) into (3.9), we deduce that

$$I(\mathfrak{P}_1) + I(\mathfrak{P}_2) \ll P^{s-k} (Z L(P) + Z^{1/2} \Xi_{\varepsilon}^{1/2})^{1-1/s} (Z L(P)^{1-3s} + Z^{1/2} \Xi_{\varepsilon}^{1/2})^{1/s}.$$

By substituting this bound into the relation (3.8), we find that

$$\begin{aligned} \psi(N)^{-1} Z &\ll L(P)^{-2} Z + L(P)^{1-1/s} \Xi_{\varepsilon}^{1/(2s)} Z^{(2s-1)/(2s)} \\ &\quad + \Xi_{\varepsilon}^{(s-1)/(2s)} Z^{(s+1)/(2s)} + \Xi_{\varepsilon}^{1/2} Z^{1/2}. \end{aligned}$$

Note that we are at liberty to suppose  $\psi(N)$  to be sufficiently small compared to  $L(P)$ , and further  $L(P)$  to be  $O(U(P)^{\varepsilon})$ . Disentangling this inequality, therefore, we conclude from (3.13) that for any positive number  $\sigma$  with  $\sigma < \sigma_1$ , one has

$$Z \ll L(P)^{2s-2} \psi(N)^{2s} \Xi_{\varepsilon} \ll P^{k-(2s-s_1)\sigma} U(P)^{\varepsilon-1}. \quad (3.14)$$

On recalling that  $P = N^{1/k}$ , the proof of Theorem 1.4 is complete.

The proof of Theorem 1.5 follows by applying a simplified variant of the above argument, as we now sketch. We make use of the notation applied in the proof of Theorem 1.4 above, fixing  $k = 3$  and  $s = 4$ . We begin by observing that the argument leading to [46, Lemma 6.1] shows on this occasion that whenever  $\mu \in (N/2, N]$ , then

$$\int_{\mathfrak{M}} \tilde{f}(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha = 2\pi \Omega_{4,3}(\boldsymbol{\lambda}, \mu/N) P + O(PL(P)^{-1}).$$

Here we note in particular that when  $k = 3$  and  $s = 4$ , the inequality (6.4) of [46] must be replaced by

$$\begin{aligned} \int_{\mathfrak{M}} f_1(\alpha) \cdots f_4(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha - \int_{\mathfrak{M}} v_1(\alpha) \cdots v_4(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha \\ \ll \sum_{n=1}^{\infty} n^{-2} \int_{-1/2}^{1/2} P^{7/2} (1 + P^3 |\alpha|)^{-1} d\alpha \ll P^{1/2} \log P. \end{aligned}$$

The latter, achieved using Lemma 2.3, suffices for the ensuing argument. We therefore see as in (3.8) that in the present situation, one has

$$I(\mathfrak{m} \cup \mathfrak{t}) \geq \frac{1}{4} P \psi(N)^{-1} Z.$$

The argument of [46, §§4 and 5], moreover, shows that

$$\int_{\mathfrak{m} \cup \mathfrak{t}} |f_1(\alpha) \cdots f_4(\alpha)|^2 K_2^*(\alpha) d\alpha \ll P^5 L(P)^{-1}.$$

Then on recalling (3.11), an application of Schwarz's inequality yields the relation

$$\begin{aligned} P \psi(N)^{-1} Z &\ll I(\mathfrak{m} \cup \mathfrak{t}) \ll \mathcal{J}^{1/2} \left( \int_{\mathfrak{m} \cup \mathfrak{t}} |f_1(\alpha) \cdots f_4(\alpha)|^2 K_2^*(\alpha) d\alpha \right)^{1/2} \\ &\ll P^{5/2} L(P)^{-1/2} Z^{1/2}, \end{aligned}$$

whence  $Z \ll P^3 \psi(N)^2 L(P)^{-1}$ . The conclusion of the theorem follows by taking  $\psi(N)$  no larger than  $L(P)^{1/4}$ .

#### 4. THE ASYMPTOTIC FORMULA FOR SMALLER EXPONENTS

In order to obtain the sharper results for  $3 \leq k \leq 5$  advertised in Theorem 1.6, we adapt the methods of Wooley [47] and Kawada and Wooley [25], [26]. The following lemma provides the appropriate analogue of [25, Lemma 6.1], which applies a method of Davenport [16] to sharpen the conclusion of [47, Lemma 2.1] in situations where the exceptional set may be relatively large.

**Lemma 4.1.** *Suppose that  $k \geq 3$ , that  $1 \leq j \leq k-2$ , and that  $\lambda$  is a non-zero real number. Let  $\mathcal{Z}$  be a subset of  $[-P^k, P^k]$  with  $\text{meas}(\mathcal{Z}) = Z$ . Then for every  $\varepsilon > 0$ , one has*

$$\int_{-\infty}^{\infty} |f(\lambda\alpha)|^{2j} H_{\eta, \mathcal{Z}}(\alpha)^2 |K_1(\alpha)| d\alpha \ll P^{2j} (P^{-1} Z + P^{-1-j/2+\varepsilon} Z^{3/2}).$$

*Proof.* We apply Weyl differencing as in the proof of [47, Lemma 2.1] to deduce that, for suitably defined intervals  $I_j(\mathbf{h}) \subseteq [1, P] \cap \mathbb{Z}$ , one has

$$|f(\lambda\alpha)|^{2j} \leq (2P)^{2^j-j-1} \sum_{\mathbf{h} \in (-P, P)^j} \sum_{x \in I_j(\mathbf{h})} e(\lambda\alpha h_1 \cdots h_j p_j(x; \mathbf{h})),$$

where  $p_j$  is a polynomial of degree  $k-j$  in  $x$ . As in the proof of Lemma 2.2, it therefore follows that

$$\int_{-\infty}^{\infty} |f(\lambda\alpha)|^{2j} H_{\eta, \mathcal{Z}}(\alpha)^2 |K_1(\alpha)| d\alpha \ll P^{2^j-j-1} \int_{\mathcal{Z}^2} \delta^{-1} \mathcal{W}(\mu, \nu) d\mu d\nu, \quad (4.1)$$

where  $\mathcal{W}(\mu, \nu)$  denotes the number of integral solutions of the inequality

$$|\lambda h_1 \cdots h_j p_j(x; \mathbf{h}) + \mu - \nu| < \delta,$$

with  $|h_i| < P$  ( $1 \leq i \leq j$ ) and  $1 \leq x \leq P$ . We let  $\mathcal{W}_0(\mu, \nu)$  denote the number of solutions  $x, \mathbf{h}$  counted by  $\mathcal{W}(\mu, \nu)$  with  $h_1 \cdots h_j = 0$ , and write  $\mathcal{W}_1(\mu, \nu)$  for the corresponding number of solutions with  $h_1 \cdots h_j \neq 0$ .

The analysis of  $\mathcal{W}_0(\mu, \nu)$  is straightforward. There are at most  $O(P^j)$  choices for  $x$  and  $\mathbf{h}$  with  $h_1 \cdots h_j = 0$ , and so  $\mathcal{W}_0(\mu, \nu) \ll P^j U_\delta(\mu - \nu)$ . Then it follows as in the proof of Lemma 2.2 that

$$\int_{\mathcal{Z}^2} \delta^{-1} \mathcal{W}_0(\mu, \nu) d\mu d\nu \ll P^j Z. \quad (4.2)$$

Turning our attention next to  $\mathcal{W}_1(\mu, \nu)$ , we denote by  $\rho(\beta, \mathbf{h})$  the number of solutions of the inequality

$$|\lambda h_1 \cdots h_j p_j(x; \mathbf{h}) + \beta| < \delta,$$

with  $1 \leq x \leq P$ . Then one has

$$\mathcal{W}_1(\mu, \nu) = \sum'_{\mathbf{h}} \rho(\mu - \nu, \mathbf{h}),$$

where we write  $\sum'$  to denote the sum over integral  $j$ -tuples  $\mathbf{h}$  with  $0 < |h_i| < P$  ( $1 \leq i \leq j$ ). We also find it convenient to write

$$\mathcal{I}_1 = \int_{\mathcal{Z}^2} \mathcal{W}_1(\mu, \nu) d\mu d\nu.$$

On applying Schwarz's inequality, followed by Cauchy's inequality, we deduce that

$$\mathcal{I}_1 \leq \left( \int_{\mathcal{Z}} d\nu \right)^{1/2} \left( \int_{\mathcal{Z}} \left| \sum'_{\mathbf{h}} \int_{\mathcal{Z}} \rho(\mu - \nu, \mathbf{h}) d\mu \right|^2 d\nu \right)^{1/2} \leq Z^{1/2} P^{j/2} \mathcal{I}_2^{1/2}, \quad (4.3)$$

where

$$\mathcal{I}_2 = \int_{\mathcal{Z}} \int_{\mathcal{Z}^2} \sum'_{\mathbf{h}} \rho(\mu_1 - \nu, \mathbf{h}) \rho(\mu_2 - \nu, \mathbf{h}) d\mu_1 d\mu_2 d\nu.$$

Write  $\mathcal{V}(\mu_1, \mu_2, \nu)$  for the number of solutions of the simultaneous inequalities

$$|\lambda h_1 \cdots h_j p_j(x_1; \mathbf{h}) + \mu_1 - \nu| < \delta \quad \text{and} \quad |\lambda h_1 \cdots h_j p_j(x_2; \mathbf{h}) + \mu_2 - \nu| < \delta,$$

with  $0 < |h_i| < P$  ( $1 \leq i \leq j$ ) and  $1 \leq x_1, x_2 \leq P$ . In addition, denote by  $\mathcal{V}_0(\mu_1, \mu_2, \nu)$  the number of solutions  $\mathbf{x}, \mathbf{h}$  counted by  $\mathcal{V}(\mu_1, \mu_2, \nu)$  in which  $p_j(x_1; \mathbf{h}) = p_j(x_2; \mathbf{h})$ , and by  $\mathcal{V}_1(\mu_1, \mu_2, \nu)$  the corresponding number of solutions with  $p_j(x_1; \mathbf{h}) \neq p_j(x_2; \mathbf{h})$ . It follows that

$$\mathcal{I}_2 = \int_{\mathcal{Z}} \int_{\mathcal{Z}^2} \left( \mathcal{V}_0(\mu_1, \mu_2, \nu) + \mathcal{V}_1(\mu_1, \mu_2, \nu) \right) d\mu_2 d\mu_1 d\nu. \quad (4.4)$$

Plainly, for each fixed choice of  $\mathbf{h}$ ,  $x_2$ , and  $\nu$ , one has

$$\int_{\mathcal{Z}} U_\delta(\lambda h_1 \cdots h_j p_j(x_2; \mathbf{h}) + \mu_2 - \nu) d\mu_2 \leq 2\delta. \quad (4.5)$$

For a given integer  $x_1$ , there are  $O(1)$  values of  $x_2$  satisfying  $p_j(x_1; \mathbf{h}) = p_j(x_2; \mathbf{h})$ , and hence by summing over  $\mathbf{h}$  and  $\mathbf{x}$  we deduce from (4.5) that

$$\int_{\mathcal{Z}} \int_{\mathcal{Z}^2} \mathcal{V}_0(\mu_1, \mu_2, \nu) d\mu_2 d\mu_1 d\nu \ll \delta \int_{\mathcal{Z}} \int_{\mathcal{Z}} \mathcal{W}_1(\mu_1, \nu) d\mu_1 d\nu = \delta \mathcal{I}_1. \quad (4.6)$$

For a solution  $\mathbf{h}, \mathbf{x}$  counted by  $\mathcal{V}_1(\mu_1, \mu_2, \nu)$ , on the other hand, we observe that the quantity  $n = h_1 \cdots h_j(p_j(x_1; \mathbf{h}) - p_j(x_2; \mathbf{h}))$  is a nonzero integer satisfying the inequality

$$|\lambda n + \mu_1 - \mu_2| < 2\delta. \quad (4.7)$$

Hence for each fixed non-zero value of  $n$ , a divisor function estimate shows that there are  $O(P^\varepsilon)$  possibilities for  $\mathbf{h}$  and  $\mathbf{x}$ . Moreover, for a given  $\mu_1$  and  $\mu_2$ , there are at most  $1 + 4\delta|\lambda|^{-1}$  integers  $n$  for which (4.7) holds. After applying the obvious analogue of (4.5) to integrate over  $\nu$ , we therefore deduce that

$$\begin{aligned} \int_{\mathcal{Z}} \int_{\mathcal{Z}^2} \mathcal{V}_1(\mu_1, \mu_2, \nu) d\mu_2 d\mu_1 d\nu &\ll \delta P^\varepsilon \int_{\mathcal{Z}^2} \sum_{n \in \mathbb{Z}} U_{2\delta}(\lambda n + \mu_1 - \mu_2) d\mu_1 d\mu_2 \\ &\ll \delta P^\varepsilon Z^2. \end{aligned} \quad (4.8)$$

Consequently, by substituting this estimate together with (4.6) into (4.4), we obtain the bound

$$\mathcal{I}_2 \ll \delta(\mathcal{I}_1 + P^\varepsilon Z^2). \quad (4.9)$$

Finally, by employing (4.9) within (4.3), we see that

$$\mathcal{I}_1 \ll (\delta P^j Z(\mathcal{I}_1 + P^\varepsilon Z^2))^{1/2},$$

whence

$$\mathcal{I}_1 \ll \delta P^j Z + \delta^{1/2} P^{j/2+\varepsilon} Z^{3/2}.$$

The lemma now follows from (4.1) and (4.2), on noting that

$$\delta^{-1/2} \ll L(P)^{1/2} \ll P^\varepsilon.$$

□

Our result for quintic forms in 31 variables instead makes use of the following analogue of [47, Lemma 3.1].

**Lemma 4.2.** *Suppose that  $k \geq 3$ , that  $2 \leq j \leq k-1$ , and that  $\lambda$  and  $\gamma$  are non-zero real numbers. Let  $\mathcal{Z}$  be a subset of  $[-P^k, P^k]$  with  $\text{meas}(\mathcal{Z}) = Z$ . Then for every  $\varepsilon > 0$ , one has*

$$\int_{-\infty}^{\infty} |f(\lambda\alpha)^{2^j} f(\gamma\alpha)^{2^{j-1}} H_{\eta, \mathcal{Z}}(\alpha)^2| K_1(\alpha) d\alpha \ll P^{2^j+2^{j-1}} (P^{-2} Z + P^{-1-j+\varepsilon} Z^2).$$

*Proof.* By proceeding as in the proof of Lemma 4.1, we find that the integral  $I$  under consideration satisfies

$$I \ll P^{2^j-j-1} \int_{\mathcal{Z}^2} \delta^{-1} \mathcal{W}(\mu, \nu) d\mu d\nu, \quad (4.10)$$

where  $\mathcal{W}(\mu, \nu)$  denotes the number of solutions of the inequality

$$\left| \lambda h_1 \cdots h_j p_j(z; \mathbf{h}) + \gamma \sum_{i=1}^{2^{j-2}} (x_i^k - y_i^k) + \mu - \nu \right| < \delta,$$

with  $|h_l| < P$  ( $1 \leq l \leq j$ ),  $1 \leq x_i, y_i \leq P$  ( $1 \leq i \leq 2^{j-2}$ ), and  $1 \leq z \leq P$ . Here, the polynomial  $p_j(z; \mathbf{h})$  that arises from the Weyl differencing process has degree  $k - j$  in  $z$ . We let  $\mathcal{W}_0(\mu, \nu)$  denote the number of solutions  $\mathbf{h}, \mathbf{x}, \mathbf{y}, z$  counted by  $\mathcal{W}(\mu, \nu)$  with  $h_1 \cdots h_j p_j(z; \mathbf{h}) = 0$ , and we write  $\mathcal{W}_1(\mu, \nu)$  for the corresponding number of solutions with  $h_1 \cdots h_j p_j(z; \mathbf{h}) \neq 0$ .

Consider first a solution  $\mathbf{h}, \mathbf{x}, \mathbf{y}, z$  counted by  $\mathcal{W}_0(\mu, \nu)$ . There are  $O(P^j)$  choices for  $\mathbf{h}$  and  $z$  satisfying  $h_1 \cdots h_j p_j(z; \mathbf{h}) = 0$ , and we therefore see that

$$\mathcal{W}_0(\mu, \nu) \ll P^j \mathcal{V}(\mu, \nu), \quad (4.11)$$

where  $\mathcal{V}(\mu, \nu)$  denotes the number of solutions of the inequality

$$\left| \gamma \sum_{i=1}^{2^{j-2}} (x_i^k - y_i^k) + \mu - \nu \right| < \delta,$$

with  $1 \leq x_i, y_i \leq P$  ( $1 \leq i \leq 2^{j-2}$ ). It follows from (2.4) that

$$\mathcal{V}(\mu, \nu) \ll \int_{-\infty}^{\infty} |f(\gamma\alpha)|^{2^{j-1}} e(\alpha(\mu - \nu)) K_+(\alpha) d\alpha,$$

where we take  $\tau = \delta$  in the definition of  $K_+$ . On substituting this estimate into (4.11), we conclude thus far that

$$\int_{\mathcal{Z}^2} \mathcal{W}_0(\mu, \nu) d\mu d\nu \ll P^j \int_{-\infty}^{\infty} |f(\gamma\alpha)|^{2^{j-1}} H_{\eta, \mathcal{Z}}(\alpha)^2 |K_+(\alpha)| d\alpha. \quad (4.12)$$

Next, by applying Weyl differencing together with the second inequality of (2.4), we find that

$$\int_{-\infty}^{\infty} |f(\gamma\alpha)|^{2^{j-1}} H_{\eta, \mathcal{Z}}(\alpha)^2 |K_+(\alpha)| d\alpha \ll P^{2^{j-1}-j} \int_{\mathcal{Z}^2} \mathcal{X}(\mu, \nu) d\mu d\nu, \quad (4.13)$$

where  $\mathcal{X}(\mu, \nu)$  denotes the number of solutions of the inequality

$$|\gamma g_1 \cdots g_{j-1} p_{j-1}(w; \mathbf{g}) + \mu - \nu| < 2\delta,$$

with  $|g_i| < P$  ( $1 \leq i \leq j-1$ ) and  $1 \leq w \leq P$ . Here the polynomial  $p_{j-1}(w; \mathbf{g})$  produced by the differencing operation is a polynomial of degree  $k - j + 1$  in  $w$ . Then by proceeding as in the arguments leading to (4.2) and (4.8) in the proof of Lemma 4.1, one finds that

$$\int_{\mathcal{Z}^2} \mathcal{X}(\mu, \nu) d\mu d\nu \ll \delta P^{j-1} Z + P^\varepsilon Z^2.$$

It therefore follows from (4.12) and (4.13) that

$$\int_{\mathcal{Z}^2} \mathcal{W}_0(\mu, \nu) d\mu d\nu \ll P^{2^{j-1}} (\delta P^{j-1} Z + P^\varepsilon Z^2). \quad (4.14)$$

Consider next a solution  $\mathbf{h}, \mathbf{x}, \mathbf{y}, z$  counted by  $\mathcal{W}_1(\mu, \nu)$ . Given any fixed one amongst the  $O(P^{2^{j-1}})$  possible choices for  $\mathbf{x}$  and  $\mathbf{y}$ , write

$$\beta(\mathbf{x}, \mathbf{y}) = \gamma \sum_{i=1}^{2^{j-2}} (x_i^k - y_i^k).$$

The number of available choices for  $\mathbf{h}$  and  $z$  is then equal to the number of solutions of the equation

$$h_1 \cdots h_j p_j(z; \mathbf{h}) = n,$$

with  $|h_i| < P$  ( $1 \leq i \leq j$ ),  $1 \leq z \leq P$ , and  $n$  a non-zero integer satisfying the inequality

$$|\lambda n + \beta(\mathbf{x}, \mathbf{y}) + \mu - \nu| < \delta.$$

It follows that a divisor function estimate once again yields the upper bound

$$\begin{aligned} \int_{\mathcal{Z}^2} \mathcal{W}_1(\mu, \nu) d\mu d\nu &\ll P^{2^{j-1}+\varepsilon} \int_{\mathcal{Z}^2} \sum_{n \in \mathbb{Z}} U_\delta(\lambda n + \beta(\mathbf{x}, \mathbf{y}) + \mu - \nu) d\mu d\nu \\ &\ll P^{2^{j-1}+\varepsilon} Z^2. \end{aligned} \quad (4.15)$$

It only remains now to combine (4.14) and (4.15) within (4.10), and we obtain the bound

$$I \ll P^{2^j-j-1} (P^{2^{j-1}} (P^{j-1} Z + \delta^{-1} P^\varepsilon Z^2)).$$

The lemma now follows on recalling again that  $\delta^{-1} \ll L(P) \ll P^\varepsilon$ .  $\square$

The conclusion of Theorem 1.6 follows from the following more general result.

**Theorem 4.3.** *Suppose that  $k \geq 3$  and that  $(s_1, \sigma_1, U)$  forms an accessible triple. Then whenever  $s \geq \frac{1}{2}s_1 + 2^{k-3}$ ,  $\sigma < \sigma_1$ , and  $\psi(N)$  grows sufficiently slowly in terms of  $s$ ,  $\sigma$ ,  $k$ ,  $\boldsymbol{\lambda}$ , and  $\tau$ , one has*

$$\tilde{Z}_{s,k}(N) \ll N^{1-1/k-(2s-s_1-2^{k-2})\sigma/k} U(N^{1/k})^{\varepsilon-1} + N^{1-2(2s-s_1-2^{k-2})\sigma/k+\varepsilon} U(N^{1/k})^{-2}.$$

Moreover, when instead  $k \geq 4$  and  $(2s-s_1-\frac{3}{8}2^k)\sigma_1 > 1$ , one has

$$\tilde{Z}_{s,k}(N) \ll N^{1-2/k-(2s-s_1-\frac{3}{8}2^k)\sigma/k} U(N^{1/k})^{-1}.$$

*Proof.* We follow the argument of the proof of Theorem 1.4 from §3, economising on details for the sake of concision. Recalling the definitions of  $\mathfrak{p}_j$ , and of  $I(\mathfrak{B})$  from (3.7), we find that  $Z$ , the measure of  $\tilde{\mathcal{Z}}_{s,k}^\tau(N; \psi; \boldsymbol{\lambda})$ , again satisfies the relation (3.8). Let  $j$  be either 1 or 2. Then an application of Hölder's inequality again yields the bound (3.9). Consider an index  $i$  with  $1 \leq i \leq s$ . Since we may suppose that  $s > k+1$ , we again obtain (3.10) as a consequence of Lemma 3.2. Next write

$$\mathcal{J}_i = \int_{-\infty}^{\infty} |f_i(\alpha)|^{2^{k-2}} |H(\alpha)|^2 K_1(\alpha) d\alpha.$$

Since we may also suppose that  $s-2^{k-3} \geq \frac{1}{2}s_1$ , an application of Schwarz's inequality in combination with Lemmata 4.1 and 3.3 leads to the bound

$$\begin{aligned} \int_{\mathfrak{n}_i} |f_i(\alpha)^s H(\alpha) K_*(\alpha)| d\alpha &\leq \mathcal{J}_i^{1/2} \mathcal{M}_{i,2s-2^{k-2}}(\mathfrak{n}_i; K_2^*)^{1/2} \\ &\ll \left( P^{2^{k-2}} (P^{-1} Z + P^{\varepsilon-k/2} Z^{3/2}) \right)^{1/2} \left( P^{2s-2^{k-2}-2k} \Xi_\varepsilon \right)^{1/2}, \end{aligned}$$

where we have written

$$\Xi_\varepsilon = P^{k-(2s-s_1-2^{k-2})(\sigma_1-\varepsilon)} U(P)^{-1}.$$

Thus we obtain the bound

$$\int_{\mathfrak{n}_i} |f_i(\alpha)^s H(\alpha) K_*(\alpha)| d\alpha \ll P^{s-k} (P^{-1}Z + P^{\varepsilon-k/2} Z^{3/2})^{1/2} \Xi_\varepsilon^{1/2}. \quad (4.16)$$

On substituting (3.10) and (4.16) into (3.9), and thence into (3.8), we deduce that

$$\begin{aligned} P^{s-k} \psi(N)^{-1} Z &\ll P^{s-k} (ZL(P) + (P^{-1/2}Z^{1/2} + P^{\varepsilon-k/4}Z^{3/4}) \Xi_\varepsilon^{1/2})^{1-1/s} \\ &\quad \times (ZL(P)^{1-3s} + (P^{-1/2}Z^{1/2} + P^{\varepsilon-k/4}Z^{3/4}) \Xi_\varepsilon^{1/2})^{1/s}. \end{aligned}$$

This inequality may be disentangled to show that

$$Z \ll L(P)^{s-1} \psi(N)^s (P^{-1/2}Z^{1/2} + P^{\varepsilon-k/4}Z^{3/4}) \Xi_\varepsilon^{1/2}.$$

We may suppose that  $\psi(N)$  is sufficiently small compared to  $L(P)$ , and that  $L(P) \ll U(P)^\varepsilon$ . Further disentangling therefore shows that for any positive number  $\sigma$  with  $\sigma < \sigma_1$ , one has

$$\begin{aligned} Z &\ll P^{-1} L(P)^{2s-2} \psi(N)^{2s} \Xi_\varepsilon + P^{\varepsilon-k} L(P)^{4s-4} \psi(N)^{4s} \Xi_\varepsilon^2 \\ &\ll P^{k-1-(2s-s_1-2^{k-2})\sigma} U(P)^{\varepsilon-1} + P^{k-2(2s-s_1-2^{k-2})\sigma+\varepsilon} U(P)^{-2}. \end{aligned}$$

On recalling that  $P = N^{1/k}$ , the proof of the first estimate of Theorem 4.3 is complete.

The second estimate of the theorem follows in like manner, once one substitutes  $\frac{3}{16}2^k$  for  $2^{k-3}$ , and Lemma 4.2 with  $j = k - 2$  for Lemma 4.1, throughout. Thus one obtains

$$\begin{aligned} P^{s-k} \psi(N)^{-1} Z &\ll P^{s-k} (ZL(P) + (P^{-1}Z^{1/2} + P^{\varepsilon-(k-1)/2}Z) \Xi_\varepsilon^{1/2})^{1-1/s} \\ &\quad \times (ZL(P)^{1-3s} + (P^{-1}Z^{1/2} + P^{\varepsilon-(k-1)/2}Z) \Xi_\varepsilon^{1/2})^{1/s}, \end{aligned}$$

where now

$$\Xi_\varepsilon = P^{k-(2s-s_1-\frac{3}{8}2^k)(\sigma_1-\varepsilon)} U(P)^{-1}.$$

We therefore conclude that whenever  $(2s - s_1 - \frac{3}{8}2^k)\sigma_1 > 1$  and  $\sigma$  is a positive number with  $\sigma < \sigma_1$ , then

$$Z \ll P^{k-2-(2s-s_1-\frac{3}{8}2^k)\sigma} U(P)^{-1}.$$

The proof of the second estimate of the theorem therefore follows again from the relation  $P = N^{1/k}$ .  $\square$

The conclusion of Theorem 1.6 follows directly from Theorem 4.3 on making use of the accessible triples recorded in the preamble to the statement of the former theorem. For all but the last column in the table, one makes use of the first estimate supplied by Theorem 4.3, and for the last column one applies the second estimate.

## 5. THE ASYMPTOTIC FORMULA IN SHORT INTERVALS

The key idea of the previous section, in which some of the excess variables were allocated to mean values involving  $|H_{\eta, \mathcal{Z}}(\alpha)|^2$ , can also be used to produce short interval results. In this section, we establish Theorem 1.7 by adjusting the analysis of Section 3 to allow for an application of Lemma 2.2 in place of Lemma 2.1.

Unless indicated otherwise, we adopt the notation of §3. Recall the hypotheses of the statement of Theorem 1.7, and suppose in particular that  $(s_1, \sigma_1, U)$  forms an accessible triple for  $k$ . Let  $\mathcal{Z} = \tilde{\mathcal{Z}}_{s+1, k}^\tau(N, M; \psi; \boldsymbol{\lambda})$ , and write

$$\lambda = \max_{1 \leq i \leq s+1} |\lambda_i| \quad \text{and} \quad c = (2(s+1)\lambda)^{-1}.$$

We begin by observing that in our proof of Theorem 1.7, it clearly suffices to show that whenever  $M \leq \lambda(k-1)(cP)^{k-1}$ , one has

$$\text{meas}(\mathcal{Z}) \ll MP^{-(2s-s_1)(\sigma_1-\varepsilon)}U(P)^{\varepsilon-1}.$$

In order to confirm this statement, we observe that if  $M > \lambda(k-1)(cP)^{k-1}$ , then we may divide the interval  $[N, N+M]$  into  $O(MP^{1-k})$  intervals of length at most  $\lambda(k-1)(cP)^{k-1}$  on which the former bound can be applied. Summing the contributions from all such intervals, the desired bound for the exceptional set follows on noting the trivial upper bound  $\lambda(k-1)(cP)^{k-1} < N^{1-1/k}$  that follows on recalling that  $P = N^{1/k}$ .

We now launch the proof of Theorem 1.7 in earnest. Suppose that  $\mu$  lies in the interval  $[P^k, P^k + M]$  and that the integers  $x_1, \dots, x_{s+1}$  satisfy

$$|\lambda_1 x_1^k + \dots + \lambda_{s+1} x_{s+1}^k - \mu| < \tau + \delta.$$

Then one has

$$\max_{1 \leq i \leq s+1} |x_i| > (2(s+1)\lambda)^{-1/k}P \geq cP.$$

Write  $\mathfrak{g}_i(\alpha)$  for  $f(\lambda_i \alpha; cP, P)$  and  $\mathfrak{h}_i(\alpha)$  for  $f(\lambda_i \alpha; 0, P)$ , in the notation of (2.1). Then we deduce from (2.3) and (2.4) that

$$\int_{-\infty}^{\infty} \prod_{i=1}^{s+1} (\mathfrak{h}_i(\alpha) - \mathfrak{g}_i(\alpha)) e(-\alpha \mu) K_{\pm}(\alpha) d\alpha = 0. \quad (5.1)$$

Next, on writing

$$\tilde{\mathfrak{h}}(\alpha) = \prod_{i=1}^{s+1} \mathfrak{h}_i(\alpha),$$

one finds that (2.3) and (2.4) yield the inequalities

$$\int_{-\infty}^{\infty} \tilde{\mathfrak{h}}(\alpha) e(-\alpha \mu) K_-(\alpha) d\alpha \leq \mathcal{N}_{s+1, k}^\tau(P; \boldsymbol{\lambda}, \mu) \leq \int_{-\infty}^{\infty} \tilde{\mathfrak{h}}(\alpha) e(-\alpha \mu) K_+(\alpha) d\alpha.$$

It therefore follows by subtracting (5.1) that

$$\sum_{\mathfrak{J} \neq \emptyset} (-1)^{|\mathfrak{J}|+1} \mathcal{R}_{\mathfrak{J}}^-(\mu; \mathbb{R}) \leq \mathcal{N}_{s+1, k}^\tau(P; \boldsymbol{\lambda}, \mu) \leq \sum_{\mathfrak{J} \neq \emptyset} (-1)^{|\mathfrak{J}|+1} \mathcal{R}_{\mathfrak{J}}^+(\mu; \mathbb{R}),$$

where the summations are over subsets  $\mathfrak{J}$  of  $\{1, \dots, s+1\}$ , and where

$$\mathcal{R}_{\mathfrak{J}}^{\pm}(\mu; \mathfrak{B}) = \int_{\mathfrak{B}} \prod_{i \in \mathfrak{J}} \mathfrak{g}_i(\alpha) \prod_{j \notin \mathfrak{J}} \mathfrak{h}_j(\alpha) e(-\alpha\mu) K_{\pm}(\alpha) d\alpha.$$

Applying the analysis leading to [46, Lemma 6.1], as in the argument following Brüdern and Wooley [15, equation (6.6)], we find that whenever  $s \geq k$ , one has

$$\sum_{\mathfrak{J} \neq \emptyset} (-1)^{|\mathfrak{J}|+1} \mathcal{R}_{\mathfrak{J}}^{\pm}(\mu; \mathfrak{M}) = 2\tau\Omega_{s,k}(\boldsymbol{\lambda}, \mu/N) P^{s+1-k} + O(P^{s+1-k} L(P)^{-1}).$$

Thus if  $\mu \in \mathcal{Z}$ , then with  $\mathcal{R}_{\mathfrak{J}}^{\mu}$  denoting either  $\mathcal{R}_{\mathfrak{J}}^+$  or  $\mathcal{R}_{\mathfrak{J}}^-$ , one has the relation

$$\sum_{\mathfrak{J} \neq \emptyset} |\mathcal{R}_{\mathfrak{J}}^{\mu}(\mu; \mathfrak{m} \cup \mathfrak{t})| > \frac{1}{2} P^{s+1-k} \psi(N)^{-1},$$

whence for some  $\mathfrak{J} \neq \emptyset$  one has

$$|\mathcal{R}_{\mathfrak{J}}^{\mu}(\mu; \mathfrak{m} \cup \mathfrak{t})| > 2^{-s-2} P^{s+1-k} \psi(N)^{-1}. \quad (5.2)$$

As in the proof of Theorem 1.4 from §3, with  $*$  equal to  $+$  or  $-$ , we denote by  $\mathcal{Z}_{\mathfrak{J}}^*$  the set of  $\mu \in \mathcal{Z}$  for which (5.2) holds with  $\mathcal{R}_{\mathfrak{J}}^{\mu} = \mathcal{R}_{\mathfrak{J}}^*$ , and we write  $Z_{\mathfrak{J}}^* = \text{meas}(\mathcal{Z}_{\mathfrak{J}}^*)$ . It follows that for some choices of  $*$  and  $\mathfrak{J}$ , one has  $Z \leq 2^{s+2} Z_{\mathfrak{J}}^*$ . We fix these choices of  $*$  and  $\mathfrak{J}$  henceforth. For each  $\mu \in \mathcal{Z}_{\mathfrak{J}}^*$ , we then determine the complex number  $\eta_{\mu}$  of modulus 1 by means of the relation

$$|\mathcal{R}_{\mathfrak{J}}^*(\mu; \mathfrak{m} \cup \mathfrak{t})| = \eta_{\mu} \mathcal{R}_{\mathfrak{J}}^*(\mu; \mathfrak{m} \cup \mathfrak{t}).$$

Integrating (5.2) over  $\mathcal{Z}_{\mathfrak{J}}^*$  gives the upper bound

$$P^{s+1-k} \psi(N)^{-1} Z \ll \int_{\mathfrak{m} \cup \mathfrak{t}} \left| \prod_{i \in \mathfrak{J}} \mathfrak{g}_i(\alpha) \prod_{j \notin \mathfrak{J}} \mathfrak{h}_j(\alpha) H(\alpha) K_*(\alpha) \right| d\alpha, \quad (5.3)$$

where in the notation introduced in (2.10), we have written  $H(\alpha) = H_{\eta, \mathcal{Z}_{\mathfrak{J}}^*}(\alpha)$ .

Suppose first that there exists an index  $i \in \mathfrak{J}$  with  $i \notin \{1, 2\}$ . In this situation, by relabelling indices if necessary, there is no loss of generality in supposing that  $s+1 \in \mathfrak{J}$ . We economise on exposition by writing

$$f_i(\alpha) = \begin{cases} \mathfrak{g}_i(\alpha), & \text{when } i \in \mathfrak{J}, \\ \mathfrak{h}_i(\alpha), & \text{when } i \notin \mathfrak{J}. \end{cases}$$

We then put

$$\tilde{f}(\alpha) = f_1(\alpha) \cdots f_s(\alpha) \quad \text{and} \quad H^{\dagger}(\alpha) = f_{s+1}(\alpha) H(\alpha),$$

and define

$$I(\mathfrak{B}) = \int_{\mathfrak{B}} |\tilde{f}(\alpha) H^{\dagger}(\alpha) K_*(\alpha)| d\alpha. \quad (5.4)$$

In this way we find that (5.3) may be rewritten in the shape

$$I(\mathfrak{p}_1) + I(\mathfrak{p}_2) \gg P^{s-k} \psi(N)^{-1} Z^{\dagger},$$

where  $Z^{\dagger} = PZ$ . Next define

$$\mathcal{J}^{\dagger} = \int_{-\infty}^{\infty} |H^{\dagger}(\alpha)|^2 K_1(\alpha) d\alpha,$$

and note that Lemma 2.2 supplies the estimate  $\mathcal{J}^\dagger \leq 2PZ = 2Z^\dagger$ . Recognising that the argument of the proof of Theorem 1.4 leading from (3.8) to (3.14) may be repeated essentially verbatim, save that  $H(\alpha)$ ,  $\mathcal{J}$  and  $Z$  are to be decorated throughout by obelisks, the estimation of  $I(\mathfrak{p}_j)$  may be completed without incident for  $j = 1, 2$ . In this way, we may conclude that whenever  $\psi(N)$  grows sufficiently slowly and  $\sigma$  is any positive number with  $\sigma < \sigma_1$ , then

$$PZ = Z^\dagger \ll P^{k-(2s-s_1)\sigma} U(P)^{\varepsilon-1},$$

and the proof of Theorem 1.7 is completed by recalling that  $P = N^{1/k}$ .

It remains to deal with the situation where  $\mathfrak{J} \subseteq \{1, 2\}$ . In this case, there is no loss of generality in supposing that  $1 \in \mathfrak{J}$ . We then put

$$\tilde{f}(\alpha) = f_2(\alpha) \cdots f_{s+1}(\alpha) \quad \text{and} \quad H^\dagger(\alpha) = f_1(\alpha)H(\alpha),$$

and define  $I(\mathfrak{B})$  again as in (5.4). In the current situation, one has

$$\sup_{\alpha \in \mathfrak{p}_1} |H^\dagger(\alpha)| \leq Z \sup_{\alpha \in \mathfrak{p}_1} |f_1(\alpha)| \ll Z^\dagger L(P)^{-3s},$$

and hence for  $1 \leq i \leq s+1$  we obtain the estimate

$$\begin{aligned} \int_{\mathfrak{N}_i \cap \mathfrak{p}_1} |f_i(\alpha)^s H^\dagger(\alpha) K_*(\alpha)| d\alpha &\ll Z^\dagger L(P)^{-3s} \mathcal{M}_{i,s}(\mathfrak{N}_i \cap \mathfrak{p}_1; |K_*|) \\ &\ll Z^\dagger P^{s-k} L(P)^{1-3s} \end{aligned}$$

as a substitute for (3.10). In all other respects, the argument outlined in the previous paragraph remains valid following a transparent relabelling of variables, and thus the conclusion of Theorem 1.7 follows even in this case.

## 6. EXCEPTIONS TO SOLUBILITY

If one is interested only in sets on which a solution to (1.1) fails to exist, rather than sets on which the asymptotic formula fails, techniques involving diminishing ranges and smooth numbers offer additional flexibility in the analysis. Diminishing ranges allow for higher-moment generalizations of Lemma 2.2, and the resulting diagonal behavior permits shorter intervals than those discussed in Theorem 1.7. Furthermore, the use of smooth numbers reduces the number of variables required to obtain best-possible mean-value estimates for the corresponding exponential sums over  $k$ th powers.

We let  $N$  be a large positive number, set  $P = N^{1/k}$ , and put

$$\lambda = \max_{1 \leq i \leq s+t} |\lambda_i|.$$

We are again free to suppose that  $\lambda_1/\lambda_2 \notin \mathbb{Q}$ . We take  $R = P^\eta$  with  $\eta > 0$  sufficiently small in terms of the ambient parameters. Suppose that  $(s_0, \sigma_0)$  forms a smooth accessible pair for  $k$ , and put  $S(P) = (\log P)^\nu$ , wherein  $\nu$  is a sufficiently small positive number. Define  $g(\alpha)$  as in (1.3) and  $f(\alpha; Q, P)$  by means of (2.1). Further, let  $c$  be a fixed positive number sufficiently large in terms of  $t, k$  and  $\boldsymbol{\lambda}$ , and write

$$P_j = c^{-j} P^{(1-1/k)^{j-1}}, \quad g_j(\alpha) = g(\lambda_j \alpha) \quad \text{and} \quad \mathfrak{f}_j(\alpha) = f(\lambda_{s+j} \alpha; P_j, 2P_j).$$

Let  $\mathcal{Z} = \mathcal{Z}_{s+t,k}^\tau(N, M; \boldsymbol{\lambda})$  denote the set of real numbers  $\mu \in [N, N+M]$  for which the inequality

$$|\lambda_1 x_1^k + \cdots + \lambda_{s+t} x_{s+t}^k - \mu| < \tau$$

has no integer solution, and write  $H_{\mathcal{Z}}(\alpha)$  for  $H_{\eta, \mathcal{Z}}(\alpha)$  when the function  $\eta$  is identically 1. The following lemma provides a natural extension of Lemma 2.2.

**Lemma 6.1.** *Let  $t$  be a positive integer, and suppose that  $M \leq P_t^{k-1}$ . In addition, let  $\mathcal{Z}$  be a subset of  $[N, N+M]$  of measure  $Z$ . Then*

$$\int_{-\infty}^{\infty} |\mathfrak{f}_1(\alpha) \cdots \mathfrak{f}_t(\alpha) H_{\mathcal{Z}}(\alpha)|^2 K_1(\alpha) d\alpha \leq 2P_1 \cdots P_t Z.$$

*Proof.* Denote by  $\mathcal{W}_m(\mu, \nu)$  the number of solutions of the inequality

$$\left| \sum_{j=m}^t \lambda_{s+j}(x_j^k - y_j^k) + \mu - \nu \right| < \delta,$$

with  $P_j < x_j, y_j \leq 2P_j$  ( $m \leq j \leq t$ ). Then as in the argument of the proof of Lemma 2.2, one finds that

$$\int_{-\infty}^{\infty} |\mathfrak{f}_1(\alpha) \cdots \mathfrak{f}_t(\alpha) H_{\mathcal{Z}}(\alpha)|^2 K_1(\alpha) d\alpha \leq \int_{\mathcal{Z}^2} \delta^{-1} \mathcal{W}_1(\mu, \nu) d\mu d\nu. \quad (6.1)$$

We show by induction that when  $\mu, \nu \in [N, N+M]$ , then one has

$$\mathcal{W}_m(\mu, \nu) \leq P_m \cdots P_t U_\delta(\mu - \nu) \quad (6.2)$$

for  $m = t, t-1, \dots, 1$ .

The case  $m = t$  follows from the argument of the proof of Lemma 2.2, so we may now suppose that  $1 \leq m < t$ , and that

$$\mathcal{W}_{m+1}(\mu, \nu) \leq P_{m+1} \cdots P_t U_\delta(\mu - \nu).$$

If  $\mathbf{x}, \mathbf{y}$  is a solution counted by  $\mathcal{W}_m(\mu, \nu)$  with  $x_m \neq y_m$ , then one has

$$|\lambda_{s+m}(x_m^k - y_m^k)| > k |\lambda_{s+m}| P_m^{k-1} = k |\lambda_{s+m}| c^{m+k} P_{m+1}^k.$$

Since we are at liberty to assume that  $c$  is sufficiently large in terms of  $t, k$  and  $\boldsymbol{\lambda}$ , and  $|\mu - \nu| \leq P_t^{k-1}$ , we arrive at the inequality

$$\left| \sum_{j=m+1}^t \lambda_{s+j}(x_j^k - y_j^k) + \mu - \nu \right| \leq 2^k \lambda t P_{m+1}^k + P_t^{k-1} < |\lambda_{s+m}(x_m^k - y_m^k)| - \delta.$$

We are therefore forced to conclude that in fact  $x_m = y_m$ , and hence the inductive hypothesis gives

$$\mathcal{W}_m(\mu, \nu) \leq P_m \mathcal{W}_{m+1}(\mu, \nu) \leq P_m (P_{m+1} \cdots P_t) U_\delta(\mu - \nu).$$

We have therefore confirmed the inductive hypothesis (6.2) for  $1 \leq m \leq t$ .

Substituting the estimate (6.2) with  $m = 1$  into (6.1), we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} |\mathfrak{f}_1(\alpha) \cdots \mathfrak{f}_t(\alpha) H_{\mathcal{Z}}(\alpha)|^2 K_1(\alpha) d\alpha &\leq \delta^{-1} P_1 \cdots P_t \int_{\mathcal{Z}} \int_{\nu-\delta}^{\nu+\delta} d\mu d\nu \\ &= 2P_1 \cdots P_t Z. \end{aligned}$$

This completes the proof of the lemma.  $\square$

Next we record an analogue of Lemma 3.1.

**Lemma 6.2.** *There exists a choice for the function  $T(P)$ , depending only on  $\lambda_1$ ,  $\lambda_2$  and  $S(P)$ , with the property that*

$$\sup_{\alpha \in \mathfrak{m}} |g_1(\alpha)g_2(\alpha)| \ll P^2 T(P)^{-5^{-k}}.$$

*Proof.* In view of our definition of the minor arcs  $\mathfrak{m}$ , the desired conclusion is immediate from [46, Lemma 8.1].  $\square$

It follows from the conclusion of Lemma 6.2 that for all  $\alpha \in \mathfrak{m}$ , one has

$$g_j(\alpha) \ll PT(P)^{-10^{-k}} \quad (6.3)$$

for at least one suffix  $j \in \{1, 2\}$ . We define  $\mathfrak{q}_j$  to be the set of real numbers  $\alpha \in \mathfrak{m}$  for which the upper bound (6.3) holds, and then put  $\mathfrak{p}_j = \mathfrak{q}_j \cup \mathfrak{t}$ , so that  $\mathfrak{m} \cup \mathfrak{t} \subseteq \mathfrak{p}_1 \cup \mathfrak{p}_2$ . In the interests of concision, we define  $v_{ij}$  just as in the discussion of §3 following (3.1).

We must introduce some additional notation before announcing an auxiliary mean value estimate. When  $1 \leq i \leq s$ , the set  $\mathfrak{B}$  is measurable, and  $K$  is integrable, we write

$$\mathcal{M}_{i,t}^*(\mathfrak{B}; K) = \int_{\mathfrak{B}} |g_i(\alpha)|^t K(\alpha) d\alpha.$$

**Lemma 6.3.** *Suppose that  $k \geq 3$ ,  $t > \max\{2k + 2, 16\}$ ,  $1 \leq i \leq s$  and  $j \in \{1, 2\}$ . Then for any fixed  $\kappa > 0$ , one has*

$$\mathcal{M}_{i,t}^*(\mathfrak{N}_i \cap \mathfrak{p}_j; |K_{\pm}|) \ll_t P^{t-k} L(P)^{1-\kappa s v_{ij}}.$$

*Proof.* We may suppose that  $t = 2k + 2 + 2\gamma$ , where  $\gamma > 0$ . We put

$$u = 2k + 2 + \gamma \quad \text{and} \quad A = 6k(8k + \gamma)/\gamma.$$

Also, we define the function  $\Psi(\alpha)$  by taking  $\Psi(\alpha) = (q + P^k|q\alpha - a|)^{-1}$ , when  $\alpha \in \mathfrak{N}(q, a) \subseteq \mathfrak{N}$ , and otherwise by putting  $\Psi(\alpha) = 0$ . Then an application of [30, Lemma 5.4] with  $\mathcal{M} = P^{3/4}$  and  $T = 2P^{1/4}$  shows that when  $\alpha \in \mathfrak{N}(q, a) \subseteq \mathfrak{N}$ , one has the estimate

$$g(\alpha) \ll P^{7/8+\varepsilon} + P(\log P)^3 q^{\varepsilon} \Psi(\alpha)^{1/(2k)}.$$

Next we observe that whenever  $q + |q\alpha - a|P^k \leq (\log P)^A$ , then one may apply [36, Lemma 8.5] to deduce that

$$g(\alpha) \ll q^{\varepsilon} P(q + |q\alpha - a|P^k)^{-1/k}.$$

By combining these estimates, therefore, we conclude that when  $\alpha \in \mathfrak{N}$  one has

$$g(\alpha) \ll P^{7/8+\varepsilon} + P\Psi(\alpha)^{2/(4k+\gamma)}.$$

Next, since  $\text{meas}(\mathfrak{N}) \ll P^{2-k}$ , it follows by a change of variable that the last estimate delivers the bound

$$\int_{\mathfrak{N}_i \cap [0,1]} |g_i(\alpha)|^u d\alpha \ll (P^{7/8+\varepsilon})^u P^{2-k} + P^{2k+\gamma} \int_{\mathfrak{N} \cap [0,1]} |g(\beta)|^2 \Psi(\beta)^{1+\gamma/(4k+\gamma)} d\beta.$$

We have  $\frac{7}{8}u + 2 < u$  whenever  $u > 16$ , and thus the hypotheses of the statement of the lemma imply that the first term on the right hand side is  $o(P^{u-k})$ . On

the other hand, a straightforward modification of the proof of [5, Lemma 2] (see Lemma 11.1 below) shows that

$$\int_{\mathfrak{N} \cap [0,1)} |g(\beta)|^2 \Psi(\beta)^{1+\gamma/(4k+\gamma)} d\beta \ll P^{2-k}.$$

We therefore conclude that

$$\int_{\mathfrak{N}_i \cap [0,1)} |g_i(\alpha)|^u d\alpha \ll P^{u-k}.$$

We recall from (2.5) that  $K_{\pm}(\alpha) \ll L(P) \min\{1, \alpha^{-2}\}$ . Then in view of the conclusion of Lemma 2.3 together with (6.3), we have on the one hand

$$\begin{aligned} \int_{\mathfrak{N}_i \cap \mathfrak{q}_j} |g_i(\alpha)|^t |K_{\pm}(\alpha)| d\alpha &\ll \left( \sup_{\alpha \in \mathfrak{q}_j} |g_i(\alpha)| \right)^{\gamma} L(P) \int_{\mathfrak{N}_i \cap [0,1)} |g_i(\alpha)|^u d\alpha \\ &\ll \left( PT(P)^{-v_{ij} 10^{-k}} \right)^{\gamma} L(P) P^{u-k} \\ &\ll P^{t-k} L(P)^{1-\kappa s v_{ij}}, \end{aligned}$$

whilst on the other

$$\begin{aligned} \int_{\mathfrak{N}_i \cap \mathfrak{t}} |g_i(\alpha)|^t |K_{\pm}(\alpha)| d\alpha &\ll g(0)^{\gamma} L(P) T(P)^{-1} \int_{\mathfrak{N}_i \cap [0,1)} |g_i(\alpha)|^u d\alpha \\ &\ll (P^{\gamma} L(P)^{-\kappa s}) P^{u-k} = P^{t-k} L(P)^{-\kappa s}. \end{aligned}$$

The conclusion of the lemma follows by combining these two upper bounds.  $\square$

**Lemma 6.4.** *Suppose that  $(s_0, \sigma_0)$  is a smooth accessible pair for  $k$ , and that  $u$  is a real number with  $u \geq s_0$ . Then for  $1 \leq i \leq s$ , there is a positive number  $\omega$  with the property that*

$$\mathcal{M}_{i,u}^*(\mathfrak{n}_i; K_2^{\pm}) \ll P^{u-k-\omega-(u-s_0)\sigma_0+\varepsilon}.$$

*Proof.* On noting (2.8), the desired estimate follows by inserting the hypothesised bounds (1.4) into the conclusion of Lemma 2.3.  $\square$

Before coming to grips with the proof of Theorem 1.1, we introduce some additional notation. Write

$$\mathcal{Z} = \mathcal{Z}_{s,k}^{\tau}(N, M; \boldsymbol{\lambda}), \quad Z = \text{meas}(\mathcal{Z}) \quad \text{and} \quad H(\alpha) = H_{\mathcal{Z}}(\alpha).$$

We define

$$\begin{aligned} \tilde{g}(\alpha) &= g_1(\alpha) \cdots g_s(\alpha), \quad \tilde{\mathfrak{f}}(\alpha) = \mathfrak{f}_1(\alpha) \cdots \mathfrak{f}_t(\alpha), \\ H^{\dagger}(\alpha) &= \tilde{\mathfrak{f}}(\alpha) H(\alpha) \quad \text{and} \quad Z^{\dagger} = \tilde{\mathfrak{f}}(0) Z. \end{aligned}$$

Also, we put

$$\mathcal{J}_1 = \int_{-\infty}^{\infty} |\tilde{\mathfrak{f}}(\alpha) H(\alpha)|^2 K_1(\alpha) d\alpha,$$

and when  $\mathfrak{B}$  is measurable, we write

$$I_1(\mathfrak{B}) = \int_{\mathfrak{B}} |\tilde{\mathfrak{f}}(\alpha) \tilde{g}(\alpha) H(\alpha) K_-(\alpha)| d\alpha.$$

As in the proof of Theorem 1.7, by subdividing the interval  $[N, N + M]$  into subintervals of length  $P_t^{k-1}$ , there is no loss of generality in supposing that  $M = P_t^{k-1}$ . If  $\mu \in \mathcal{Z}$ , then it follows from (2.3) that

$$\int_{-\infty}^{\infty} \tilde{f}(\alpha) \tilde{g}(\alpha) e(-\alpha\mu) K_-(\alpha) d\alpha = 0. \quad (6.4)$$

When  $s \geq k+1$ , moreover, the analysis leading to [46, Lemma 9.4] is easily modified to confirm that, uniformly in  $\mu \in [N, N + M]$ , one has

$$\int_{\mathfrak{M}} \tilde{f}(\alpha) \tilde{g}(\alpha) e(-\alpha\mu) K_-(\alpha) d\alpha \gg \tau \tilde{f}(0) P^{s-k}.$$

Then by subtracting (6.4) and integrating over all  $\mu \in \mathcal{Z}$ , we deduce that

$$I_1(\mathfrak{p}_1) + I_1(\mathfrak{p}_2) \gg \tau P^{s-k} Z^\dagger. \quad (6.5)$$

Let  $j$  be either 1 or 2. Then an application of Hölder's inequality reveals that

$$I_1(\mathfrak{p}_j) \leq \prod_{i=1}^s \left( \int_{\mathfrak{p}_j} |g_i(\alpha)^s H^\dagger(\alpha) K_-(\alpha)| d\alpha \right)^{1/s}. \quad (6.6)$$

Consider an index  $i$  with  $1 \leq i \leq s$ . In view of the hypotheses of the statement of Theorem 1.1, we may suppose that  $s > \max\{2k+2, 16\}$ , and thus we deduce from Lemma 6.3 that

$$\begin{aligned} \int_{\mathfrak{N}_i \cap \mathfrak{p}_j} |g_i(\alpha)^s H^\dagger(\alpha) K_-(\alpha)| d\alpha &\leq H^\dagger(0) \mathcal{M}_{i,s}^*(\mathfrak{N}_i \cap \mathfrak{p}_j; |K_-|) \\ &\ll Z^\dagger P^{s-k} L(P)^{1-3sv_{ij}}. \end{aligned} \quad (6.7)$$

We may also suppose that  $s \geq \frac{1}{2}s_0$ , and so an application of Schwarz's inequality in combination with Lemmata 6.1 and 6.4 leads to the bound

$$\begin{aligned} \int_{\mathfrak{n}_i} |g_i(\alpha)^s H^\dagger(\alpha) K_-(\alpha)| d\alpha &\leq \mathcal{J}_1^{1/2} \mathcal{M}_{i,2s}^*(\mathfrak{n}_i; K_2^-)^{1/2} \\ &\ll (Z^\dagger)^{1/2} (P^{2s-2k} \Xi_\varepsilon)^{1/2}, \end{aligned} \quad (6.8)$$

where we write

$$\Xi_\varepsilon = P^{k-\omega-(2s-s_0)\sigma_0+\varepsilon}. \quad (6.9)$$

On substituting (6.7) and (6.8) into (6.6), we deduce that

$$\begin{aligned} I_1(\mathfrak{p}_1) + I_1(\mathfrak{p}_2) &\ll P^{s-k} (Z^\dagger L(P) + (Z^\dagger)^{1/2} \Xi_\varepsilon^{1/2})^{1-1/s} \\ &\quad \times (Z^\dagger L(P)^{1-3s} + (Z^\dagger)^{1/2} \Xi_\varepsilon^{1/2})^{1/s}. \end{aligned}$$

Substituting this bound into (6.5), we find that

$$P^{s-k} Z^\dagger \ll P^{s-k} L(P)^{-3} (Z^\dagger L(P) + P^\varepsilon \Xi_\varepsilon^{1/2} (Z^\dagger)^{1/2}).$$

Note that

$$\sum_{j=1}^t (1 - 1/k)^{j-1} = k - k(1 - 1/k)^t,$$

and hence  $\tilde{f}(0) \gg P^k P_t^{1-k} = P^k M^{-1}$ . Then on disentangling the last inequality, we conclude from (6.9) that there exists a positive number  $\Delta$  for which

$$P^k M^{-1} Z \ll Z^\dagger \ll P^\varepsilon \Xi_\varepsilon \ll P^{k-k\Delta-(2s-s_0)\sigma_0},$$

whence  $Z \ll M P^{-k\Delta-(2s-s_0)\sigma_0}$ . On recalling that  $P = N^{1/k}$ , the proof of Theorem 1.1 is complete.

## 7. SOLUBILITY FOR SMALLER EXPONENTS

For smaller values of  $k$  we have available the option of applying a slightly different argument using some generating functions on a complete interval. This enables us to take advantage of the fact that Weyl's inequality is superior to the currently available values for  $\sigma$  in (1.4) when  $k \leq 6$ . We also gain an advantage in our major arc treatment through the use of Lemma 3.2 instead of Lemma 6.3. A full account of the details and ramifications of Theorem 1.2 would demand much more space than seems warranted for the present paper. We therefore aim for a concise exposition in which some of the less significant details are sketched rather than fully explained.

We begin by indicating how to establish the conclusion of Theorem 1.2 for  $k \geq 7$ . Recall the integers  $s_0(k)$ ,  $u_0(k)$  and  $\sigma(k)$  recorded in Table 1. Then by reference to the tables of permissible exponents in [38, §§9-22], one finds from the last numerical value in each table that there is a positive integer  $t < \frac{1}{2}s_0$  together with positive numbers  $\lambda_t$ ,  $\sigma$  and  $\omega$ , such that

$$\begin{aligned} \int_{\mathfrak{n} \cap [0,1)} |g(\alpha)|^{s_0} d\alpha &\leq \left( \sup_{\alpha \in \mathfrak{n}} |g(\alpha)| \right)^{s_0-2t} \int_0^1 |g(\alpha)|^{2t} d\alpha \\ &\ll (P^{1-\sigma+\varepsilon})^{s_0-2t} P^{\lambda_t+\varepsilon} \ll P^{s_0-k-\omega}. \end{aligned}$$

For each integer  $k$  with  $7 \leq k \leq 20$ , therefore, one finds that the exponent pair  $(s_0(k), \sigma(k))$  defined in Table 1, forms a smooth accessible pair for  $k$ . The analogous conclusion holds for larger values of  $k$  by virtue of the results contained in [39] and [41]. The conclusion of Theorem 1.2 therefore follows at once from Theorem 1.1 for  $k \geq 7$ , and it remains only to consider the exponents  $k$  with  $4 \leq k \leq 6$ .

When  $4 \leq k \leq 6$ , define the integers  $u = u(k)$ ,  $v = v(k)$  and  $w = w(k)$  as in the table below.

$k$	4	5	6
$u(k)$	5	8	12
$v(k)$	12	20	26
$w(k)$	8	12	16

Table 3: Parameters for the proof of Theorem 1.2

We note for future reference that for each exponent  $k$ , one has

$$z(k) = (w - u)/(1 - u/v) > k + 1.$$

On considering the underlying Diophantine equations, the methods of [34] and [37] show that for each  $k$  one has

$$\int_0^1 |g(\alpha)|^v d\alpha \ll P^{v-k}. \quad (7.1)$$

This argument is made explicit for  $k = 4$  in [34, Lemma 5.2] and for  $k = 5$  in [37, Lemma 7.3]. The reader should not experience any difficulty in extracting the analogous conclusion for  $k = 6$  in like manner. Finally, we note that from the tables of exponents in [12] and [37], one has

$$\int_0^1 |g(\alpha)|^{2u} d\alpha \ll P^{2u-k+\theta+\varepsilon}, \quad (7.2)$$

where  $\theta = \theta_{u,k}$  is given by

$$\theta_{5,4} = 0.213431, \quad \theta_{8,5} = 0.077363, \quad \text{and} \quad \theta_{12,6} = 0.$$

We suppose that  $s \geq w = \frac{1}{2}(s_0 + u_0)$ , and employ the notation introduced in §6, modifying the definition of  $\tilde{g}(\alpha)$  by putting

$$\tilde{g}(\alpha) = g_1(\alpha) \cdots g_u(\alpha) f_{u+1}(\alpha) \cdots f_s(\alpha).$$

As in the proof of Theorem 1.7, by subdividing the interval  $[N, N + M]$  into subintervals of length  $P_t^{k-1}$ , there is no loss of generality in supposing that  $M = P_t^{k-1}$ . When  $s \geq k + 1$ , moreover, the analysis leading to [46, Lemma 9.4] is again easily modified to confirm that, uniformly in  $\mu \in [N, N + M]$ , one has

$$\int_{\mathfrak{M}} \tilde{\mathfrak{f}}(\alpha) \tilde{g}(\alpha) e(-\alpha\mu) K_-(\alpha) d\alpha \gg \tau \tilde{\mathfrak{f}}(0) P^{s-k}.$$

Then we deduce as in the argument of the proof of Theorem 1.1 that the lower bound (6.5) remains valid in the present circumstances.

Let  $j$  be either 1 or 2. Then an application of Hölder's inequality reveals that

$$I_1(\mathfrak{p}_j) \leq \prod_{i=1}^u \prod_{m=u+1}^s \left( \int_{\mathfrak{p}_j} |g_i(\alpha)^u f_m(\alpha)^{s-u} H^\dagger(\alpha) K_-(\alpha)| d\alpha \right)^{1/(u(s-u))}. \quad (7.3)$$

Consider indices  $i$  and  $m$  with  $1 \leq i \leq u < m \leq s$ . An application of Hölder's inequality yields the bound

$$\begin{aligned} & \int_{\mathfrak{N}_m \cap \mathfrak{p}_j} |g_i(\alpha)^u f_m(\alpha)^{s-u} H^\dagger(\alpha) K_-(\alpha)| d\alpha \\ & \leq H^\dagger(0) \mathcal{M}_{i,v}^*(\mathbb{R}; |K_-|)^{u/v} \mathcal{M}_{m,y}(\mathfrak{N}_m \cap \mathfrak{p}_j; |K_-|)^{1-u/v}, \end{aligned}$$

where

$$y = (s - u)/(1 - u/v) \geq (w - u)/(1 - u/v) = z.$$

Since we may suppose that  $z > k + 1$ , we deduce from (7.1) via Lemma 2.3 in combination with Lemma 3.2 that

$$\int_{\mathfrak{N}_m \cap \mathfrak{p}_j} |g_i(\alpha)^u f_m(\alpha)^{s-u} H^\dagger(\alpha) K_-(\alpha)| d\alpha \ll Z^\dagger P^{s-k} L(P)^{1-3sv_{ij}}. \quad (7.4)$$

Next, by applying the mean value estimate (7.2) together with Lemma 2.3, we find that

$$\mathcal{M}_{i,2u}^*(\mathbb{R}; K_2^-) \ll P^{2u-k+\theta+\varepsilon}.$$

Then on recalling Weyl's inequality (see [35, Lemma 2.4]), it follows from an application of Schwarz's inequality in combination with Lemma 6.1 that

$$\begin{aligned} \int_{\mathfrak{n}_m} |g_i(\alpha)^u f_m(\alpha)^{s-u} H^\dagger(\alpha) K_-(\alpha)| d\alpha &\leq \left( \sup_{\alpha \in \mathfrak{n}_m} |f_m(\alpha)| \right)^{s-u} \mathcal{J}_1^{1/2} \mathcal{M}_{i,2u}^*(\mathbb{R}; K_2^-)^{1/2} \\ &\ll (Z^\dagger)^{1/2} (P^{1-\sigma+\varepsilon})^{s-u} (P^{2u-k+\theta+\varepsilon})^{1/2}. \end{aligned}$$

Since  $\sigma = 2^{1-k}$ , we see that there exists a positive number  $\Delta$  with the property that

$$\int_{\mathfrak{n}_m} |g_i(\alpha)^u f_m(\alpha)^{s-u} H^\dagger(\alpha) K_-(\alpha)| d\alpha \ll (Z^\dagger)^{1/2} (P^{2s-2k} \Xi)^{1/2}, \quad (7.5)$$

where we write

$$\Xi = P^{k-k\Delta-(2s-s_0)\sigma}.$$

On substituting (7.4) and (7.5) into (7.3) and from there into (6.5), we deduce that

$$P^{s-k} Z^\dagger \ll P^{s-k} (Z^\dagger L(P) + (Z^\dagger)^{1/2} \Xi^{1/2})^{1-1/s} (Z^\dagger L(P)^{1-3s} + (Z^\dagger)^{1/2} \Xi^{1/2})^{1/s}.$$

Then on disentangling the last inequality, we conclude just as before that

$$P^k M^{-1} Z \ll Z^\dagger \ll P^{k-k\Delta-(2s-s_0)\sigma},$$

whence  $Z \ll M P^{-k\Delta-(2s-s_0)\sigma}$ . On recalling that  $P = N^{1/k}$ , the proof of the main conclusion of Theorem 1.2 is complete.

We now turn to a brief sketch of the proof of the estimate  $Z_{s+t,k}(N, M) \ll MN^{-\Delta}$ , for a sufficiently small  $\Delta > 0$ , valid for  $s \geq \frac{1}{2}s_0$ . Here we must obtain greater control of the behaviour of the generating functions  $g_i(\alpha)$  on the major arcs  $\mathfrak{N}_i$ . This we achieve by means of two modifications to the above argument. First we replace the underlying sequence  $\mathcal{A}(P, R)$  by the related sequence

$$\mathcal{C}(P, R) = \{lm : 1 \leq l \leq \sqrt{R}, 1 \leq m \leq P/\sqrt{R}, p|m \Rightarrow \sqrt{R} < p \leq R\}.$$

We also replace the major arc  $\mathfrak{N}(q, a)$  by the set of real numbers  $\alpha$  for which  $|q\alpha - a| \leq \sqrt{R}P^{-k}$ , and then write  $\mathfrak{N}$  for the union of the intervals  $\mathfrak{N}(q, a)$  over all co-prime integers  $a$  and  $q$  with  $1 \leq q \leq \sqrt{R}$ . We then put  $\mathfrak{n} = \mathbb{R} \setminus \mathfrak{N}$ . As the reader will find in the papers [13] and [14], with such a modification, the generating function  $g(\alpha)$  may be analysed on  $\mathfrak{N}$  essentially as precisely as the generating function  $f(\alpha)$ . Thus one finds that the conclusion of Lemma 6.3 holds with the hypothesis on  $t$  weakened to the condition  $t > k + 1$ . On the other hand, the conclusion of Lemma 6.4 now holds only with  $\sigma_0$  replaced by a positive number depending on  $\eta$ , which we recall defines  $R$  by means of the relation  $R = P^\eta$ .

With the modifications described in the previous paragraph, the argument of §6 employed in the proof of Theorem 1.1 applies without further modification to show that when  $s \geq \max\{k + 1, \frac{1}{2}s_0\}$ , then for some positive number  $\Delta$  one has

$$Z_{s+t,k}(N, M) \ll MN^{-\Delta}.$$

This completes our sketch of the proof of the remaining part of Theorem 1.2, the details, though not difficult, being lengthy to record in full.

## 8. CUBIC FORMS

In this section we outline the proof of Theorem 1.3. Although in principle one has only to apply the methods of §7 together with Lemma 4.1 to establish the desired bounds for  $Z_{s,3}(N)$  ( $s = 5, 6$ ), the small number of available variables leads to complications in handling the contributions from the major arcs. We note that the bound  $Z_{4,3}(N) = o(N)$  is a consequence of Theorem 1.5, which we have already established. Thus it suffices to consider  $Z_{s,3}(N)$  for  $s = 5$  and  $6$ .

We begin by establishing an auxiliary mean value estimate. We suppose throughout that  $k = 3$  and  $P = N^{1/3}$ . Also, we let  $\mathfrak{V}(q, a)$  denote the set of real numbers  $\alpha \in [0, 1)$  with  $|q\alpha - a| \leq P^{-9/4}$ , and write  $\mathfrak{V}$  for the union of the sets  $\mathfrak{V}(q, a)$  with  $0 \leq a \leq q \leq P^{3/4}$  and  $(a, q) = 1$ . We then put  $\mathfrak{v} = [0, 1) \setminus \mathfrak{V}$ .

**Lemma 8.1.** *Suppose that  $\mathcal{Z} \subseteq [0, N]$  has measure  $Z$ , and  $|\eta_\mu| = 1$  for all  $\mu \in \mathcal{Z}$ . Then whenever  $\lambda$  is a non-zero real number, one has*

$$\int_{-\infty}^{\infty} |g(\lambda\alpha)^4 H_{\eta, \mathcal{Z}}(\alpha)^2| K_1(\alpha) d\alpha \ll P^3 (\log P)^{2+\varepsilon} Z + P (\log P)^\varepsilon Z^2.$$

*Proof.* As in the proof of Lemma 2.1, the mean value under consideration may be written as

$$I = \int_{\mathcal{Z}^2} \bar{\eta}_\mu \eta_\nu \int_{-\infty}^{\infty} |g(\lambda\alpha)|^4 e(\alpha(\mu - \nu)) K_1(\alpha) d\alpha d\mu d\nu.$$

The relation (2.9) shows that

$$I \leq \int_{\mathcal{Z}^2} \delta^{-1} \mathcal{W}(\mu, \nu) d\mu d\nu, \quad (8.1)$$

where

$$\mathcal{W}(\mu, \nu) = \sum_{1 \leq x_1, \dots, x_4 \leq P} \max\{0, 1 - \delta^{-1} |\lambda(x_1^3 + x_2^3 - x_3^3 - x_4^3) + \mu - \nu|\}.$$

We observe that on writing  $R(h)$  for the number of integral solutions of the equation

$$x_1^3 + x_2^3 - x_3^3 - x_4^3 = h,$$

with  $1 \leq x_i \leq P$  ( $1 \leq i \leq 4$ ), then one has

$$\mathcal{W}(\mu, \nu) = \sum_{h \in \mathbb{Z}} R(h) \max\{0, 1 - \delta^{-1} |\lambda h + \mu - \nu|\}. \quad (8.2)$$

When  $\mathfrak{B} \subseteq [0, 1)$  is measurable, write

$$R(h; \mathfrak{B}) = \int_{\mathfrak{B}} |f(\alpha)|^4 e(-h\alpha) d\alpha.$$

We analyse  $R(h)$  by means of the Hardy-Littlewood method, and in this way we deduce from (8.1) and (8.2) that

$$\int_{-\infty}^{\infty} |g(\lambda\alpha)^4 H_{\eta, \mathcal{Z}}(\alpha)^2| K_1(\alpha) d\alpha \ll T(\mathfrak{V}) + T(\mathfrak{v}), \quad (8.3)$$

where we write

$$T(\mathfrak{B}) = \delta^{-1} \int_{\mathcal{Z}^2} \sum_{h \in \mathbb{Z}} R(h; \mathfrak{B}) \max\{0, 1 - \delta^{-1}|\lambda h + \mu - \nu|\} d\mu d\nu.$$

We begin by analysing the major arc contribution. Although there is no convenient source in the literature, the reader will experience no difficulty in applying the methods of [35, §§4.3 and 4.4] to confirm that

$$R(h; \mathfrak{V}) \ll P(\log P)^\varepsilon \quad (h \neq 0), \quad \text{and} \quad R(0; \mathfrak{V}) \ll P^{1+\varepsilon}.$$

We note that a precise form of the first of these bounds may be found in [24, equation (1.3)]. Equipped with these bounds, we deduce that

$$\begin{aligned} T(\mathfrak{V}) &\ll \delta^{-1} P^{1+\varepsilon} \int_{\mathcal{Z}^2} U_\delta(\mu - \nu) d\mu d\nu \\ &\quad + \delta^{-1} P(\log P)^\varepsilon \int_{\mathcal{Z}^2} \sum_{h \in \mathbb{Z}} U_\delta(\lambda h + \mu - \nu) d\mu d\nu \\ &\ll \delta^{-1} P^{1+\varepsilon} \int_{\mathcal{Z}} \int_{\nu-\delta}^{\nu+\delta} d\mu d\nu + \delta^{-1} P(\log P)^\varepsilon \int_{\mathcal{Z}^2} d\mu d\nu. \end{aligned}$$

Consequently, one has

$$T(\mathfrak{V}) \ll P^{1+\varepsilon} Z + P(\log P)^\varepsilon Z^2. \quad (8.4)$$

Next, we recall that an enhanced version of Weyl's inequality is available from the argument underlying the proof of [31, Lemma 1], on applying bounds of Hall and Tenenbaum [21] for Hooley's  $\Delta$ -function in place of Vaughan's application of Hooley [23]. Thus one finds that

$$\sup_{\alpha \in \mathfrak{v}} |f(\alpha)| \ll P^{3/4} (\log P)^{1/4+\varepsilon}. \quad (8.5)$$

Observe next that

$$T(\mathfrak{v}) = \sum_{|h| \leqslant (N+\delta)/\lambda} R(h; \mathfrak{v}) \int_{-\infty}^{\infty} |H_{\mathcal{Z}}(\beta)|^2 e(\lambda h \beta) K_1(\beta) d\beta.$$

But equipped with the estimate (8.5), we deduce that

$$\begin{aligned} \sum_{|h| \leqslant (N+\delta)/\lambda} R(h; \mathfrak{v}) e(\lambda h \beta) &= \int_{\mathfrak{v}} |f(\alpha)|^4 \sum_{|h| \leqslant (N+\delta)/\lambda} e(h(\lambda \beta - \alpha)) d\alpha \\ &\ll \int_{\mathfrak{v}} |f(\alpha)|^4 \min\{(N+\delta)/\lambda, \|\alpha - \lambda \beta\|^{-1}\} d\alpha \\ &\ll P^3 (\log P)^{1+\varepsilon} \int_0^1 \min\{(N+\delta)/\lambda, \|\alpha - \lambda \beta\|^{-1}\} d\alpha. \end{aligned}$$

We therefore obtain the bound

$$\sum_{|h| \leqslant (N+\delta)/\lambda} R(h; \mathfrak{v}) e(\lambda h \beta) \ll P^3 (\log P)^{2+\varepsilon},$$

and hence we conclude from Lemma 2.1 that

$$T(\mathfrak{v}) \ll P^3(\log P)^{2+\varepsilon} \int_{-\infty}^{\infty} |H_{\mathcal{Z}}(\beta)|^2 K_1(\beta) d\beta \ll P^3(\log P)^{2+\varepsilon} Z. \quad (8.6)$$

The proof of the lemma is completed by substituting the estimates (8.4) and (8.6) into (8.3).  $\square$

We also require a variant of Lemma 6.3 providing a reasonably sharp bound with relatively few variables. We must first introduce some auxiliary sets of major arcs. When  $1 \leq i \leq 5$ , we write  $\mathfrak{V}_i$  for the set of real numbers  $\alpha \in \mathfrak{m} \cup \mathfrak{t}$  for which  $\lambda_i \alpha \in \mathfrak{V} \pmod{1}$ , and  $\mathfrak{v}_i = (\mathfrak{m} \cup \mathfrak{t}) \setminus \mathfrak{V}_i$ . In addition, we define  $\mathfrak{W}(q, a)$  to be the set of  $\alpha \in [0, 1)$  such that  $|q\alpha - a| \leq (\log P)^{100} P^{-3}$ , and take  $\mathfrak{W}$  to be the union of the sets  $\mathfrak{W}(q, a)$  with  $0 \leq a \leq q \leq (\log P)^{100}$  and  $(a, q) = 1$ . We then put  $\mathfrak{w} = [0, 1) \setminus \mathfrak{W}$ . Also, when  $1 \leq i \leq 5$ , we write  $\mathfrak{W}_i$  for the set of real numbers  $\alpha \in \mathfrak{m} \cup \mathfrak{t}$  for which  $\lambda_i \alpha \in \mathfrak{W} \pmod{1}$ , and  $\mathfrak{w}_i = (\mathfrak{m} \cup \mathfrak{t}) \setminus \mathfrak{W}_i$ .

We pause to record an auxiliary lemma.

**Lemma 8.2.** *Suppose that  $t > 5$  and  $1 \leq i \leq s$ . Then one has*

$$\mathcal{M}_{i,t}^*(\mathfrak{W}_i; |K_{\pm}|) \ll_t P^{t-3} L(P).$$

*Proof.* We may suppose that  $t = 2 + (3 + \gamma)(1 + \gamma)$ , where  $\gamma > 0$ . We define the function  $\Psi(\alpha)$  by taking  $\Psi(\alpha) = (q + P^3|q\alpha - a|)^{-1}$  when  $\alpha \in \mathfrak{W}(q, a) \subseteq \mathfrak{W}$ , and otherwise by putting  $\Psi(\alpha) = 0$ . One may apply [36, Lemma 8.5] to deduce that when  $\alpha \in \mathfrak{W}(q, a) \subseteq \mathfrak{W}$ , then

$$g(\alpha) \ll q^{\varepsilon} P(q + |q\alpha - a|P^3)^{-1/3}.$$

Then when  $\alpha \in \mathfrak{W}$  one has

$$g(\alpha) \ll P\Psi(\alpha)^{1/(3+\gamma)}.$$

It follows by a change of variable that the last estimate delivers the bound

$$\int_{\mathfrak{W}_i \cap [0,1)} |g_i(\alpha)|^t d\alpha \ll P^{t-2} \int_{\mathfrak{W}} |g(\beta)|^2 \Psi(\beta)^{1+\gamma} d\beta.$$

A straightforward modification of the proof of [5, Lemma 2] (see Lemma 11.1 below) shows that

$$\int_{\mathfrak{W}} |g(\beta)|^2 \Psi(\beta)^{1+\gamma} d\beta \ll P^{-1}.$$

We therefore arrive at the upper bound

$$\int_{\mathfrak{W}_i \cap [0,1)} |g_i(\alpha)|^t d\alpha \ll P^{t-3}.$$

The conclusion of the lemma consequently follows from Lemma 2.3.  $\square$

Now we initiate the proof of Theorem 1.3. We write  $\mathcal{Z} = \mathcal{Z}_{s,3}^{\tau}(N; \boldsymbol{\lambda})$ , and put  $Z_{s,3}(N) = \text{meas}(\mathcal{Z})$ . We seek first to bound  $Z = Z_{5,3}(N)$ , and fix  $s = 5$  in this first case. Write

$$\tilde{g}(\alpha) = f_1(\alpha)f_2(\alpha)g_3(\alpha)g_4(\alpha)g_5(\alpha),$$

and then put

$$I(\mathfrak{B}) = \int_{\mathfrak{B}} |\tilde{g}(\alpha)H(\alpha)K_-(\alpha)| d\alpha, \quad (8.7)$$

in which we have written  $H(\alpha) = H_{\mathcal{Z}}(\alpha)$ . We also write

$$\mathcal{J}_m = \int_{-\infty}^{\infty} |g_m(\alpha)^4 H(\alpha)^2| K_1(\alpha) d\alpha.$$

The analysis leading to [46, Lemma 9.4] is easily modified to confirm that, uniformly in  $\mu \in (N/2, N]$ , one has

$$\int_{\mathfrak{M}} \tilde{g}(\alpha) e(-\alpha\mu) K_-(\alpha) d\alpha \gg \tau P^2.$$

Moreover, for each  $\mu \in \mathcal{Z}$  one has

$$\int_{-\infty}^{\infty} \tilde{g}(\alpha) e(-\alpha\mu) K_-(\alpha) d\alpha = 0.$$

Then by subtracting and integrating over  $\mu \in \mathcal{Z}$ , we deduce as before that

$$I(\mathfrak{p}_1) + I(\mathfrak{p}_2) \gg P^2 Z. \quad (8.8)$$

Let  $j$  be either 1 or 2. Then an application of Hölder's inequality to (8.7) reveals that

$$I(\mathfrak{p}_j) \leq \prod_{i=1}^2 \prod_{m=1}^3 \left( \int_{\mathfrak{p}_j} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \right)^{1/6}.$$

Consider indices  $i$  and  $m$  with  $i \in \{1, 2\}$  and  $m \in \{3, 4, 5\}$ . By applying Hölder's inequality, we find that

$$\begin{aligned} & \int_{\mathfrak{V}_i \cap \mathfrak{W}_m \cap \mathfrak{p}_j} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \\ & \leq H(0) \mathcal{M}_{i,22/5}(\mathfrak{V}_i \cap \mathfrak{p}_j; |K_-|)^{5/11} \mathcal{M}_{m,11/2}^*(\mathfrak{W}_m; |K_-|)^{6/11}. \end{aligned}$$

From Lemmata 3.2 and 8.2, therefore, we deduce that

$$\begin{aligned} & \int_{\mathfrak{V}_i \cap \mathfrak{W}_m \cap \mathfrak{p}_j} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \\ & \leq H(0) (P^{7/5} L(P)^{1-9sv_{ij}})^{5/11} (P^{5/2} L(P))^{6/11} \\ & \ll Z P^2 L(P)^{1-3sv_{ij}}. \end{aligned} \quad (8.9)$$

Next, an application of Hölder's inequality reveals that

$$\begin{aligned} & \int_{\mathfrak{V}_i \cap \mathfrak{w}_m} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \\ & \leq \left( \sup_{\alpha \in \mathfrak{w}_m} |g_m(\alpha)| \right)^{7/11} \mathcal{M}_{i,22/5}(\mathfrak{V}_i \cap \mathfrak{p}_j; K_2^-)^{5/11} \mathcal{J}_m^{1/2} \mathcal{M}_{m,8}^*(\mathbb{R}; K_2^-)^{1/22}. \end{aligned}$$

On the one hand, we observe that as a consequence of [36, Lemmata 7.2 and 8.5] together with [34, Theorem 1.8], one has

$$\sup_{\alpha \in \mathfrak{w}_m} |g_m(\alpha)| \ll P (\log P)^{-30}.$$

On the other hand, by considering the underlying Diophantine equations, it follows from Vaughan [31, Theorem 2] that

$$\int_0^1 |g(\beta)|^8 d\beta \ll P^5,$$

whence from Lemma 2.3 we obtain

$$\mathcal{M}_{m,8}^*(\mathbb{R}; K_2^-) \ll P^5. \quad (8.10)$$

We therefore deduce from Lemmata 3.2 and 8.1 that

$$\begin{aligned} \int_{\mathfrak{V}_i \cap \mathfrak{w}_m} & |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \\ & \ll (P(\log P)^{-30})^{7/11} (P^{7/5})^{5/11} (P^5)^{1/22} \\ & \quad \times (P^3 (\log P)^{2+\varepsilon} Z + P(\log P)^\varepsilon Z^2)^{1/2}, \end{aligned}$$

whence

$$\int_{\mathfrak{V}_i \cap \mathfrak{w}_m} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \ll P^3 Z^{1/2} + P^2 (\log P)^{-2} Z. \quad (8.11)$$

Finally, an application of Schwarz's inequality in combination with Lemma 2.3 delivers the bound

$$\int_{\mathfrak{v}_i} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \ll \left( \sup_{\alpha \in \mathfrak{v}} |f_i(\alpha)| \right)^2 \left( \int_0^1 |g(\alpha)|^6 d\alpha \right)^{1/2} \mathcal{J}^{1/2},$$

where  $\mathcal{J}$  is defined as in (3.11). Define  $\theta$  by means of the relation  $\theta^{-1} = 852 + 16\sqrt{2833}$ . Then, by utilising [43, Theorem 1.2] together with (8.5) and Lemma 2.1, we find that

$$\begin{aligned} \int_{\mathfrak{v}_i} & |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \ll (P^{3/4+\varepsilon})^2 (P^{13/4-\theta+\varepsilon})^{1/2} Z^{1/2} \\ & \ll P^{25/8-\theta/2+3\varepsilon} Z^{1/2}. \end{aligned} \quad (8.12)$$

By combining (8.9), (8.11) and (8.12), we reach the bound

$$\int_{\mathfrak{p}_j} |f_i(\alpha)^2 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \ll P^2 L(P)^{1-3sv_{ij}} Z + P^{25/8-\theta/2+\varepsilon} Z^{1/2}.$$

We therefore deduce that

$$\begin{aligned} I(\mathfrak{p}_1) + I(\mathfrak{p}_2) & \ll \left( Z P^2 L(P) + P^{25/8-\theta/2+\varepsilon} Z^{1/2} \right)^{1/2} \\ & \quad \times \left( Z P^2 L(P)^{-6} + P^{25/8-\theta/2+\varepsilon} Z^{1/2} \right)^{1/2}, \end{aligned}$$

so that we may conclude from (8.8) that

$$Z P^2 \ll Z P^2 L(P)^{-1} + P^{25/8-\theta/2+\varepsilon} Z^{1/2} + Z^{3/4} P^{41/16-\theta/4+\varepsilon}.$$

Disentangling this inequality, we arrive at the bound

$$Z \ll P^{9/4-\theta+\varepsilon} \ll N^{3/4-\theta/3+\varepsilon}.$$

This completes the proof of the estimate for  $Z_{5,3}(N)$  asserted in Theorem 1.3.

We now turn to the bound for  $Z_{6,3}(N)$  recorded in Theorem 1.3. On this occasion we write

$$\tilde{g}(\alpha) = f_1(\alpha)f_2(\alpha)f_3(\alpha)g_4(\alpha)g_5(\alpha)g_6(\alpha),$$

and then define  $I(\mathfrak{B})$  as in (8.7). We also write

$$\mathcal{J}_i = \int_{-\infty}^{\infty} |f_i(\alpha)|^2 H(\alpha)^2 |K_-(\alpha)| d\alpha.$$

The analysis leading to [46, Lemma 9.4] is easily modified to confirm that, uniformly in  $\mu \in (N/2, N]$ , one has

$$\int_{\mathfrak{M}} \tilde{g}(\alpha) e(-\alpha\mu) K_-(\alpha) d\alpha \gg \tau P^3.$$

Then we deduce as before that

$$I(\mathfrak{p}_1) + I(\mathfrak{p}_2) \gg P^3 Z. \quad (8.13)$$

Let  $j$  be either 1 or 2. Then an application of Hölder's inequality reveals that

$$I(\mathfrak{p}_j) \leq \prod_{i=1}^3 \prod_{m=1}^3 \left( \int_{\mathfrak{p}_j} |f_i(\alpha)|^3 g_m(\alpha)^3 H(\alpha) K_-(\alpha) d\alpha \right)^{1/9}. \quad (8.14)$$

Consider indices  $i$  and  $m$  with  $i \in \{1, 2, 3\}$  and  $m \in \{4, 5, 6\}$ . By applying Hölder's inequality, we find that

$$\begin{aligned} & \int_{\mathfrak{V}_i \cap \mathfrak{p}_j} |f_i(\alpha)|^3 g_m(\alpha)^3 H(\alpha) K_-(\alpha) d\alpha \\ & \leq H(0) \mathcal{M}_{i,24/5}(\mathfrak{V}_i \cap \mathfrak{p}_j; |K_-|)^{5/8} \mathcal{M}_{m,8}^*(\mathbb{R}; |K_-|)^{3/8}. \end{aligned}$$

By employing an argument akin to that delivering the bound (8.10), one finds that  $\mathcal{M}_{m,8}^*(\mathbb{R}; K^-) \ll P^5 L(P)$ . Then from Lemma 3.2, we deduce that

$$\begin{aligned} & \int_{\mathfrak{V}_i \cap \mathfrak{p}_j} |f_i(\alpha)|^3 g_m(\alpha)^3 H(\alpha) K_-(\alpha) d\alpha \\ & \leq H(0) (P^{9/5} L(P)^{1-9sv_{ij}})^{5/8} (P^5 L(P))^{3/8} \ll Z P^3 L(P)^{1-3sv_{ij}}. \end{aligned} \quad (8.15)$$

Next, an application of Schwarz's inequality in combination with Lemma 2.3 reveals that

$$\int_{\mathfrak{v}_i} |f_i(\alpha)|^3 g_m(\alpha)^3 H(\alpha) K_-(\alpha) d\alpha \ll \left( \sup_{\alpha \in \mathfrak{v}_i} |f_i(\alpha)| \right)^2 \left( \int_0^1 |g_m(\alpha)|^6 d\alpha \right)^{1/2} \mathcal{J}_i^{1/2}.$$

Then as a consequence of [43, Theorem 1.2] together with Lemma 4.1 and (8.5), we find that

$$\begin{aligned} & \int_{\mathfrak{v}_i} |f_i(\alpha)|^3 g_m(\alpha)^3 H(\alpha) K_-(\alpha) d\alpha \\ & \ll (P^{3/4+\varepsilon})^2 (P^{13/4-\theta+\varepsilon})^{1/2} (PZ + P^{1/2+\varepsilon} Z^{3/2})^{1/2} \\ & \ll P^{29/8-\theta/2+3\varepsilon} Z^{1/2} + P^{27/8-\theta/2+3\varepsilon} Z^{3/4}. \end{aligned} \quad (8.16)$$

Combining (8.15) and (8.16), we obtain the bound

$$\begin{aligned} & \int_{\mathfrak{p}_j} |f_i(\alpha)^3 g_m(\alpha)^3 H(\alpha) K_-(\alpha)| d\alpha \\ & \ll P^3 L(P)^{1-3sv_{ij}} Z + P^{29/8-\theta/2+\varepsilon} Z^{1/2} + P^{27/8-\theta/2+\varepsilon} Z^{3/4}, \end{aligned}$$

and hence by (8.14) we have

$$\begin{aligned} I(\mathfrak{p}_1) + I(\mathfrak{p}_2) & \ll \left( P^3 L(P) Z + P^{29/8-\theta/2+\varepsilon} Z^{1/2} + P^{27/8-\theta/2+\varepsilon} Z^{3/4} \right)^{2/3} \\ & \times \left( Z P^3 L(P)^{-6} + P^{29/8-\theta/2+\varepsilon} Z^{1/2} + P^{27/8-\theta/2+\varepsilon} Z^{3/4} \right)^{1/3}. \end{aligned}$$

We may therefore conclude from (8.13) that

$$Z P^3 \ll Z P^3 L(P)^{-1} + P^{29/8-\theta/2+\varepsilon} Z^{1/2} + P^{27/8-\theta/2+\varepsilon} Z^{3/4}.$$

Disentangling this inequality, we arrive at the bound

$$Z \ll P^{5/4-\theta+\varepsilon} + P^{3/2-2\theta+\varepsilon} \ll N^{1/2-2\theta/3+\varepsilon}.$$

This completes the proof of the estimate for  $Z_{6,3}(N)$  asserted in Theorem 1.3.

## 9. LOWER BOUND THEOREMS

The proof of Theorem 1.8 is a very simple argument based on use of the kernel

$$K(\alpha) = \left( \frac{\sin \pi \alpha \tau}{\pi \alpha} \right)^2, \quad (9.1)$$

which is obtained by setting  $\delta = \tau$  in the definition of  $\delta^2 K_1(\alpha)$  given by (2.7). Let  $\mathcal{Y} = \mathcal{Y}_{s,k}^\tau(N; \boldsymbol{\lambda})$ , and write  $Y = \text{meas}(\mathcal{Y})$  and  $H(\alpha) = H_{\mathcal{Y}}(\alpha)$ . We have

$$\int_{-\infty}^{\infty} g_1(\alpha) \cdots g_s(\alpha) H(\alpha) K(\alpha) d\alpha = \sum_{\mathbf{x} \in \mathcal{A}(P,R)^s} \int_{\mathcal{Y}} \widehat{K}(\lambda_1 x_1^k + \cdots + \lambda_s x_s^k - \mu) d\mu,$$

and it follows from (2.9) that  $\widehat{K}(\beta) = \max\{0, \tau - |\beta|\}$ . Hence for each  $\mathbf{x}$  there is an interval  $[F(\mathbf{x}) - \tau/2, F(\mathbf{x}) + \tau/2]$  contained in  $\mathcal{Y}$  on which  $\widehat{K}(F(\mathbf{x}) - \mu) \geq \tau/2$ , and it follows that

$$\int_{-\infty}^{\infty} g_1(\alpha) \cdots g_s(\alpha) H(\alpha) K(\alpha) d\alpha \gg \tau^2 P^s.$$

Applying Hölder's inequality, we therefore obtain

$$\tau^2 P^s \ll \left( \int_{-\infty}^{\infty} |H(\alpha)|^2 K(\alpha) d\alpha \right)^{1/2} \prod_{j=1}^s \left( \int_{-\infty}^{\infty} |g_j(\alpha)|^{2s} K(\alpha) d\alpha \right)^{1/(2s)}.$$

Thus if  $\Delta = \Delta_{s,k}$  is an admissible exponent, then we deduce from (1.9), (2.7), Lemma 2.1, and Lemma 2.3 that

$$\tau^2 P^s \ll \tau Y^{1/2} (P^{2s-k+\Delta+\varepsilon})^{1/2},$$

and hence  $Y \gg \tau^2 P^{k-\Delta-\varepsilon}$ . In particular, we deduce from [43, Theorem 1.2] that the exponent  $\Delta_{3,3} = (\sqrt{2833} - 43)/41 = 0.249413\dots$  is admissible, and this completes the proof of Theorem 1.8.

## 10. LINEAR COMBINATIONS OF TWO PRIMES

In view of the analysis of [28] and the pattern of argument established in the previous sections, we can be somewhat brief in our proof of Theorem 1.9. Here we employ the weighted exponential sums

$$h_i(\alpha) = \sum_{p \leq X} (\log p) e(\lambda_i p \alpha),$$

and we write  $\mathcal{Z} = \mathcal{Z}^*(X; \boldsymbol{\lambda}, \tau) \cap (X/2, X]$  and  $Z = \text{meas}(\mathcal{Z})$ . We let  $\mathfrak{M}$  denote the set of  $\alpha$  satisfying  $|\alpha| \leq (\log X)^A X^{-1}$ , for a suitable constant  $A > 0$ , and we define the boundary between the minor and trivial arcs using a suitable function  $T(X)$ , with the property that

$$\sup_{\alpha \in \mathfrak{m}} |h_1(\alpha)h_2(\alpha)| \ll X^2 L(X)^{-1}, \quad (10.1)$$

wherein the function  $L(X) \ll T(X)$  tends to infinity sufficiently slowly. The existence of such functions  $T$  and  $L$  follows from [28] (see Lemma 2 and the discussion in §3). We let  $K(\alpha)$  be as in (9.1) and write  $H(\alpha) = H_{\mathcal{Z}}(\alpha)$ . If  $\mu \in \mathcal{Z}$ , then one has

$$\int_{-\infty}^{\infty} h_1(\alpha)h_2(\alpha)e(-\alpha\mu)K(\alpha) d\alpha = 0,$$

and the analysis of [28, §4] yields

$$\int_{\mathfrak{M}} h_1(\alpha)h_2(\alpha)e(-\alpha\mu)K(\alpha) d\alpha \gg \tau^2 \mu.$$

Hence on integrating over  $\mathcal{Z}$  and applying Schwarz's inequality, we find that

$$\tau^2 X Z \ll \mathcal{I}_1^{1/2} \mathcal{I}_2^{1/2}, \quad (10.2)$$

where

$$\mathcal{I}_1 = \int_{-\infty}^{\infty} |H(\alpha)|^2 K(\alpha) d\alpha \quad \text{and} \quad \mathcal{I}_2 = \int_{\mathfrak{M} \cup \mathfrak{t}} |h_1(\alpha)h_2(\alpha)|^2 K(\alpha) d\alpha.$$

By Lemma 2.1 we have  $\mathcal{I}_1 \ll \tau^2 Z$ . Furthermore, by applying [28, Lemma 3] and the trivial bound  $|z_1 z_2| \leq |z_1|^2 + |z_2|^2$  together with Lemma 2.3, we obtain

$$\mathcal{I}_2 \ll X^2 \sup_{\alpha \in \mathfrak{m}} |h_1(\alpha)h_2(\alpha)|^{1/2} + X^3 T(X)^{-1}.$$

We therefore deduce from (10.1) and (10.2) that

$$\tau X Z \ll Z^{1/2} X^{3/2} L(X)^{-1/4},$$

and hence that  $Z \ll \tau^{-2} X L(X)^{-1/2}$ . Theorem 1.9 now follows on summing over dyadic intervals.

## 11. APPENDIX: A VARIANT OF BRÜDERN'S PRUNING LEMMA

It is convenient to have available a sharp version of Brüdern's pruning lemma which avoids the loss of  $\varepsilon$ -powers of the basic parameter. Although technically a straightforward modification of [5, Lemma 2], we supply details here in order to provide a complete exposition. Our argument is motivated by the proof of [8, Lemma 3.3].

We begin with some notation. Let  $k$  be a natural number with  $k \geq 2$ , let  $N$  be a large real number, and put  $P = N^{1/k}$ . We suppose that  $\mathcal{A} \subseteq [1, P] \cap \mathbb{Z}$ , and we write

$$F(\alpha) = \sum_{x \in \mathcal{A}} e(\alpha x^k).$$

Finally, we define the multiplicative function  $w_k(q)$  by taking

$$w_k(p^{uk+v}) = \begin{cases} kp^{-u-1/2}, & \text{when } u \geq 0 \text{ and } v = 1, \\ p^{-u-1}, & \text{when } u \geq 0 \text{ and } 2 \leq v \leq k. \end{cases}$$

Then according to [32, Lemma 3], whenever  $a \in \mathbb{Z}$  and  $q \in \mathbb{N}$  satisfy  $(a, q) = 1$ , the exponential sum

$$S_k(q, a) = \sum_{r=1}^q e(ar^k/q)$$

satisfies the bound  $q^{-1}S_k(q, a) \ll w_k(q)$ .

**Lemma 11.1.** *Let  $Q$  be a real number with  $Q \leq P$ . When  $a \in \mathbb{Z}$  and  $q \in \mathbb{N}$  satisfy  $1 \leq a \leq q \leq Q$  and  $(a, q) = 1$ , let  $\mathcal{M}(q, a)$  denote an interval contained in  $[a/q - \frac{1}{2}, a/q + \frac{1}{2}]$ , and assume that the sets  $\mathcal{M}(q, a)$  are disjoint. Write  $\mathcal{M}$  for the union of the sets  $\mathcal{M}(q, a)$ . Also, let  $\gamma$  be a positive number, and let  $G : \mathcal{M} \rightarrow \mathbb{C}$  be a function satisfying*

$$G(\alpha) \ll (q + N|q\alpha - a|)^{-1-\gamma} \quad \text{for } \alpha \in \mathcal{M}(q, a).$$

Then

$$\int_{\mathcal{M}} G(\alpha)|F(\alpha)|^2 d\alpha \ll_{\gamma} P^2 N^{-1}.$$

*Proof.* One has

$$\int_{\mathcal{M}} G(\alpha)|F(\alpha)|^2 d\alpha \ll \sum_{1 \leq q \leq Q} q^{-1-\gamma} \int_{-\infty}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^q \frac{|F(a/q + \beta)|^2}{(1 + N|\beta|)^{1+\gamma}} d\beta. \quad (11.1)$$

Write  $c_q(h)$  for Ramanujan's sum, which we define by

$$c_q(h) = \sum_{\substack{a=1 \\ (a,q)=1}}^q e(ah/q).$$

Then it follows that

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q |F(a/q + \beta)|^2 = \sum_{x,y \in \mathcal{A}} c_q(x^k - y^k) e(\beta(x^k - y^k)).$$

The estimate  $|c_q(h)| \leq (q, h)$  therefore conveys us from (11.1) to the bound

$$\int_{\mathcal{M}} G(\alpha) |F(\alpha)|^2 d\alpha \ll_{\gamma} N^{-1} \sum_{1 \leq q \leq Q} q^{-1-\gamma} \sum_{1 \leq x, y \leq P} (q, x^k - y^k). \quad (11.2)$$

Write  $\rho(d)$  for the number of solutions of the congruence  $x^k \equiv y^k \pmod{d}$  with  $1 \leq x, y \leq d$ . Then, by sorting  $x$  and  $y$  into residue classes modulo  $d$ , we find that whenever  $q \leq P$  one has

$$\begin{aligned} \sum_{1 \leq x, y \leq P} (q, x^k - y^k) &\leq \sum_{d|q} d \cdot \text{card}\{1 \leq x, y \leq P : x^k \equiv y^k \pmod{d}\} \\ &\leq \sum_{d|q} (P/d + 1)^2 d \rho(d) \ll P^2 \sum_{d|q} \rho(d)/d. \end{aligned} \quad (11.3)$$

But  $\rho(d)$  is a multiplicative function of  $d$ , and when  $d$  is a prime power  $p^h$  with  $h \geq 1$ , one has

$$\begin{aligned} \rho(p^h) &= p^{-h} \sum_{b=1}^{p^h} |S_k(p^h, b)|^2 = p^{-h} \sum_{l=0}^h \sum_{\substack{c=1 \\ (c,p)=1}}^{p^l} (p^{h-l} |S_k(p^l, c)|)^2 \\ &\leq p^h \sum_{l=0}^h p^l w_k(p^l)^2 \leq k^2(h+1)p^{2h} w_k(p^h)^2. \end{aligned}$$

Since  $\rho(d)$  is a multiplicative function of  $d$ , then  $\sum_{d|q} \rho(d)/d$  must be a multiplicative function of  $q$ . When  $q$  is a prime power  $p^m$  with  $m \geq 1$ , therefore, one finds that

$$\sum_{d|q} \rho(d)/d \leq \sum_{h=0}^m k^2(h+1)p^h w_k(p^h)^2 \leq k^4(m+1)^2 p^m w_k(p^m)^2.$$

Consequently, we deduce from (11.3) that

$$\begin{aligned} q^{-1-\gamma} \sum_{1 \leq x, y \leq P} (q, x^k - y^k) &\ll P^2 \prod_{\substack{p^m \parallel q \\ m \geq 1}} ((p^m)^{-1-\gamma} k^4(m+1)^2 p^m w_k(p^m)^2) \\ &\leq P^2 \prod_{\substack{p^m \parallel q \\ m \geq 1}} (k^6(m+1)^2 p^{-1-m\gamma}). \end{aligned} \quad (11.4)$$

Next, on substituting (11.4) into (11.2), we deduce that

$$\begin{aligned} \int_{\mathcal{M}} G(\alpha) |F(\alpha)|^2 d\alpha &\ll_{\gamma} P^2 N^{-1} \sum_{q=1}^{\infty} \prod_{\substack{p^m \parallel q \\ m \geq 1}} (k^6(m+1)^2 p^{-1-m\gamma}) \\ &= P^2 N^{-1} \prod_p \left( 1 + p^{-1} \sum_{m=1}^{\infty} k^6(m+1)^2 p^{-m\gamma} \right) \\ &\leq P^2 N^{-1} \prod_p (1 + A p^{-1-\gamma}), \end{aligned}$$

for some positive number  $A$  depending at most on  $k$ . Thus we may conclude that

$$\begin{aligned} \int_{\mathcal{M}} G(\alpha) |F(\alpha)|^2 d\alpha &\ll_{\gamma} P^2 N^{-1} \prod_p (1 - p^{-1-\gamma})^{-A} \\ &\leq P^2 N^{-1} \zeta(1 + \gamma)^A \ll_{\gamma} P^2 N^{-1}. \end{aligned}$$

This completes the proof of the lemma.  $\square$

## REFERENCES

- [1] R. C. Baker, *Diophantine inequalities*, London Mathematical Society Monographs, New Series, vol. 1, Oxford University Press, Oxford, 1986.
- [2] V. Bentkus and F. Götzte, *Lattice point problems and distribution of values of quadratic forms*, Annals of Math. (2) **150** (1999), 977–1027.
- [3] K. D. Boklan, *A reduction technique in Waring's problem, I*, Acta Arith. **65** (1993), 147–161.
- [4] K. D. Boklan, *The asymptotic formula in Waring's problem*, Mathematika **41** (1994), 329–347.
- [5] J. Brüdern, *A problem in additive number theory*, Math. Proc. Cambridge Philos. Soc. **103** (1988), 27–33.
- [6] J. Brüdern, *On Waring's problem for cubes*, Math. Proc. Cambridge Philos. Soc. **109** (1991), 229–256.
- [7] J. Brüdern, R. J. Cook and A. Perelli, *The values of binary linear forms at prime arguments*, Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), London Math. Soc. Lecture Note Ser. 237, Cambridge University Press, Cambridge, 1997, pp. 87–100.
- [8] J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, I: Waring's problem for cubes*, Ann. Sci. École Norm. Sup. (4) **34** (2001), 471–501.
- [9] J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, III: asymptotic formulae*, Acta Arith. **100** (2001), 267–289.
- [10] J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, VIII: Diophantine inequalities in review*, Number Theory. Dreaming in Dreams, Proceedings of the 5th China-Japan Seminar, Higashi-Osaka, 2008 (T. Aoki, et al., eds.), World Scientific, 2009, pp. 20–79.
- [11] J. Brüdern, K. Kawada and T. D. Wooley, *Annexe to the gallery: an addendum to “Additive representation in thin sequences, VIII: Diophantine inequalities in review”*, submitted, 6pp.
- [12] J. Brüdern and T. D. Wooley, *On Waring's problem: two cubes and seven biquadrates*, Tsukuba J. Math. **24** (2000), 387–417.
- [13] J. Brüdern and T. D. Wooley, *On Waring's problem for cubes and smooth Weyl sums*, Proc. London Math. Soc. (3) **82** (2001), 89–109.
- [14] J. Brüdern and T. D. Wooley, *On Waring's problem: three cubes and a sixth power*, Nagoya Math. J. **163** (2001), 13–53.
- [15] J. Brüdern and T. D. Wooley, *Additive representation in short intervals, I: Waring's problem for cubes*, Compositio Math. **140** (2004), 1197–1220.
- [16] H. Davenport, *On sums of positive integral kth powers*, Amer. J. Math. **64** (1942), 189–198.
- [17] H. Davenport and H. Heilbronn, *On indefinite quadratic forms in five variables*, J. London Math. Soc. **21** (1946), 185–193.
- [18] K. B. Ford, *New estimates for mean values of Weyl sums*, Internat. Math. Res. Notices (1995), no. 3, 155–171.
- [19] D. E. Freeman, *Asymptotic lower bounds for Diophantine inequalities*, Mathematika **47** (2000), 127–159.
- [20] D. E. Freeman, *Asymptotic lower bounds and formulas for Diophantine inequalities*, Number Theory for the Millennium (Urbana, IL, 2000) (M. A. Bennett et. al., ed.), vol. 2, 2002, pp. 57–74.

- [21] R. Hall and G. Tenenbaum, *Divisors*, Cambridge University Press, Cambridge, 1988.
- [22] D. R. Heath-Brown, *Weyl's inequality, Hua's inequality, and Waring's problem*, J. London Math. Soc. (2) **38** (1988), 216–230.
- [23] C. Hooley, *On a new technique and its applications to the theory of numbers*, Proc. London Math. Soc. (3) **38** (1979), 115–151.
- [24] K. Kawada, *On the sum of four cubes*, Mathematika **43** (1996), 323–348.
- [25] K. Kawada and T. D. Wooley, *Relations between exceptional sets for additive problems*, J. London Math. Soc. (2) **82** (2010), 437–458.
- [26] K. Kawada and T. D. Wooley, *Davenport's method and slim exceptional sets: the asymptotic formulae in Waring's problem*, Mathematika **56** (2010), 305–321.
- [27] A. V. Kumchev, *On the Waring-Goldbach problem: exceptional sets for sums of cubes and higher powers*, Canad. J. Math. **57** (2005), 298–327.
- [28] S. T. Parsell, *Irrational linear forms in prime variables*, J. Number Theory **97** (2002), 144–156.
- [29] S. T. Parsell, *On simultaneous diagonal inequalities, III*, Quart. J. Math. **53** (2002), 347–363.
- [30] S. T. Parsell and T. D. Wooley, *On pairs of diagonal quintic forms*, Compositio Math. **131** (2002), 61–96.
- [31] R. C. Vaughan, *On Waring's problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170.
- [32] R. C. Vaughan, *On Waring's problem for smaller exponents*, Proc. London Math. Soc. (3) **52** (1986), 445–463.
- [33] R. C. Vaughan, *On Waring's problem for smaller exponents, II*, Mathematika **33** (1986), 6–22.
- [34] R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71.
- [35] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd ed., Cambridge University Press, Cambridge, 1997.
- [36] R. C. Vaughan and T. D. Wooley, *On Waring's problem: some refinements*, Proc. London Math. Soc. (3) **63** (1991), 35–68.
- [37] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240.
- [38] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem, IV: higher powers*, Acta Arith. **94** (2000), 203–285.
- [39] T. D. Wooley, *Large improvements in Waring's problem*, Annals of Math. (2) **135** (1992), 131–164.
- [40] T. D. Wooley, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399.
- [41] T. D. Wooley, *New estimates for smooth Weyl sums*, J. London Math. Soc. (2) **51** (1995), 1–13.
- [42] T. D. Wooley, *Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour*, Invent. Math. **122** (1995), 421–451.
- [43] T. D. Wooley, *Sums of three cubes*, Mathematika **47** (2000), 53–61.
- [44] T. D. Wooley, *Slim exceptional sets for sums of cubes*, Canad. J. Math. **54** (2002), 417–448.
- [45] T. D. Wooley, *Slim exceptional sets for sums of four squares*, Proc. London Math. Soc. (3) **85** (2002), 1–21.
- [46] T. D. Wooley, *On Diophantine inequalities: Freeman's asymptotic formulae*, Proceedings of the session in analytic number theory and Diophantine equations (Bonn, January–June, 2002) (D. R. Heath-Brown and B. Z. Moroz, eds.), no. 360, Bonner Mathematische Schriften, 2003.
- [47] T. D. Wooley, *Slim exceptional sets and the asymptotic formula in Waring's problem*, Math. Proc. Cambridge Philos. Soc. **134** (2003), 193–206.
- [48] T. D. Wooley, *The asymptotic formula in Waring's problem*, Internat. Math. Res. Notices (2012), no. 7, 1485–1504.
- [49] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing*, Annals of Math. (2) **175** (2012), 1575–1627.

- [50] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing, II*, Duke Math. J. (accepted, to appear).

STP: DEPARTMENT OF MATHEMATICS, WEST CHESTER UNIVERSITY, 25 UNIVERSITY AVE., WEST CHESTER, PA 19383, U.S.A.

*E-mail address:* sparsell@wcupa.edu

TDW: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRISTOL, UNIVERSITY WALK, CLIFTON, BRISTOL BS8 1TW, UNITED KINGDOM

*E-mail address:* matdw@bristol.ac.uk